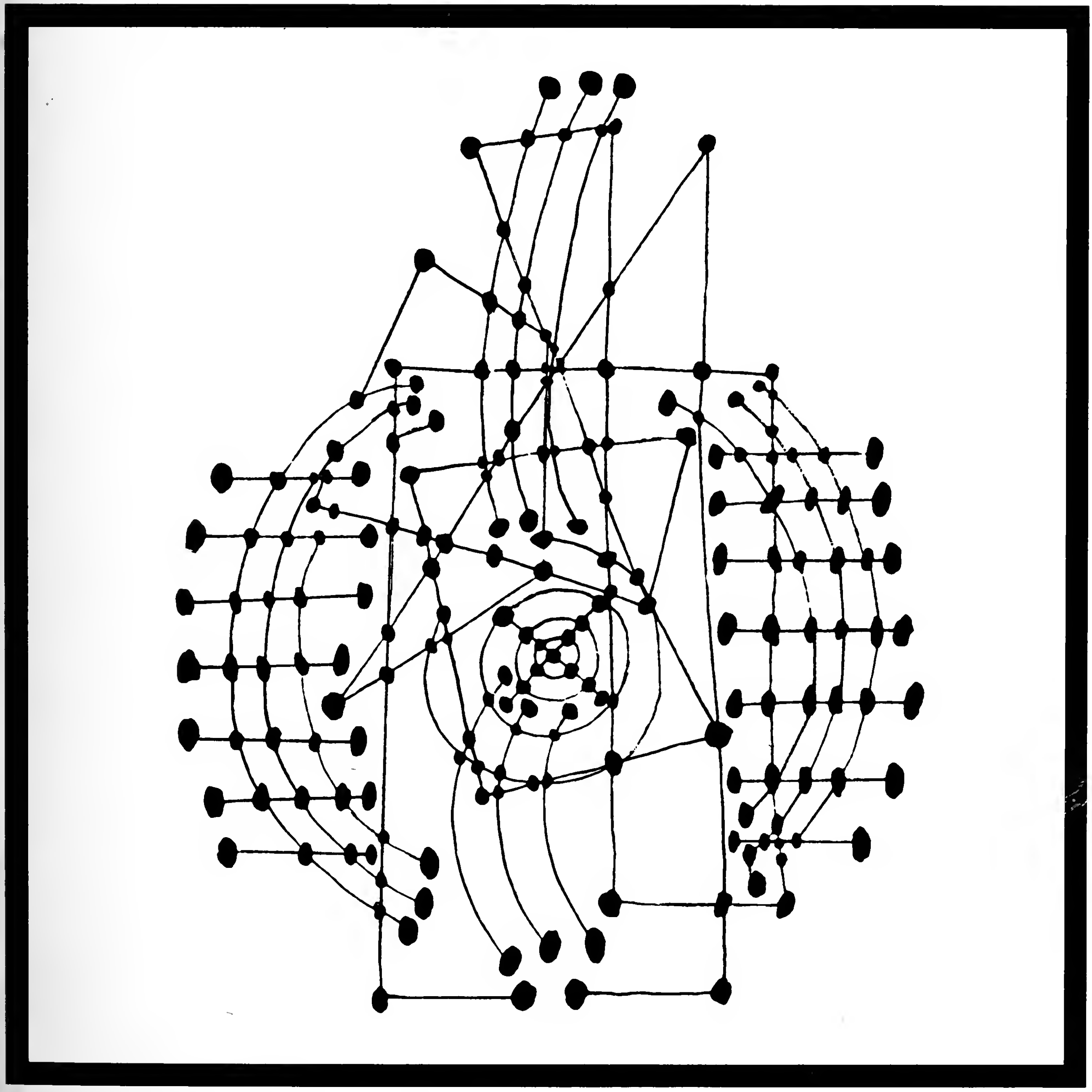
The Project Physics Course

Reader 6

The Nucleus



The Project Physics Course

Reader

UNIT 6 The Nucleus

A Component of the
Project Physics Course

Published by
HOLT, RINEHART and WINSTON, Inc.
New York, Toronto

This publication is one of the many instructional materials developed for the Project Physics Course. These materials include Texts, Handbooks, Teacher Resource Books, Readers, Programmed Instruction Booklets, Film Loops, Transparencies, 16mm films and laboratory equipment. Development of the course has profited from the help of many colleagues listed in the text units.

Directors of Harvard Project Physics

Gerald Holton, Department of Physics, Harvard University

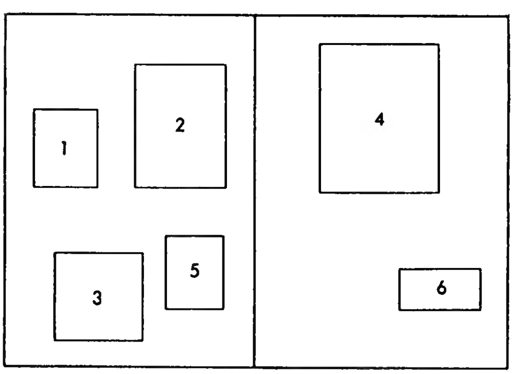
F. James Rutherford, Capuchino High School, San Bruno, California, and Harvard University

Fletcher G. Watson, Harvard Graduate School of Education

SBN 03-084563-7
1234 039 98765432

Picture Credits

Cover picture: Ink drawing by Pablo Picasso from *Le Chel-d'oeuvre inconnu,* by Honoré de Balzac, Ambroise Vollard, Paris, 1931.



Picture Credits for frontispiece.

(1) Photograph by Glen J. Pearcy.
(2) *Jeune fille au corsage rouge lisant* by Jean Baptiste Camille Corot. Painting. Collection Bührle, Zurich.
(3) Harvard Project Physics staff photo.
(4) *Femme lisant* by Georges Seurat. Conté crayon drawing. Collection C. F. Stoop, London.
(5) *Portrait of Pierre Reverdy* by Pablo Picasso. Etching. Museum of Modern Art, N.Y.C.
(6) *Lecture au lit* by Paul Klee. Drawing. Paul Klee Foundation, Museum of Fine Arts, Berne.

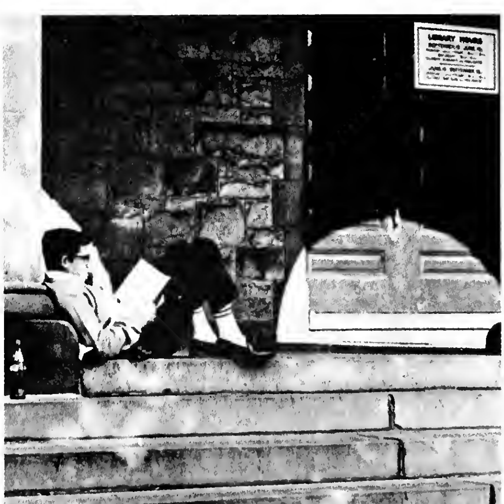
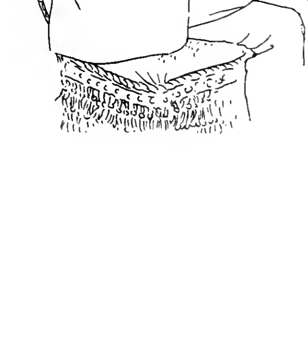
Sources and Acknowledgments
Project Physics Reader 6

1. *Rutherford* from *Variety of Men,* pp. 3–30, by C. P. Snow, copyright © 1967 by C. P. Snow. Reprinted with the permission of Charles Scribner's Sons.
2. *The Nature of the Alpha Particle from Radioactive Substances* by E. Rutherford and T. Royds from the *Philosophical Magazine,* Chapter 6, Volume 17, 1909, pp. 281–286. Reproduced by permission of Taylor and Francis Ltd., London.
3. *Some Personal Notes on the Search for the Neutron* by Sir James Chadwick from *Actes du Xeme Congres International d'Histoire des Sciences,* Hermann, Paris and from *Ithaca,* 26 VIII–2 IX 1962. Reprinted with permission.
4. *Anti-Protons* by O. Chamberlain, E. Segré, C. Wiegand, and T. Ypsilantis from *Nature,* Volume 177, January 7, 1956, pages 11–12. Reprinted with permission.
5. *The Tracks of Nuclear Particles* by Herman Yagoda from *Scientific American,* May 1956, copyright © 1956 by Scientific American, Inc. Reprinted with permission. All rights reserved. Available separately at 25¢ each as Offprint No. 252 from W. H. Freeman and Company, 660 Market Street, San Francisco, California 94104.
6. *The Spark Chamber* by Gerard K. O'Neill from *Scientific American,* August 1962, copyright © 1962 by Scientific American, Inc. Reprinted with permission. All rights reserved. Available separately at 25¢ each as Offprint No. 282 from W. H. Freeman and Company, 660 Market Street, San Francisco, California 94104.
7. *The Evolution of the Cyclotron* by E. O. Lawrence. (Nobel Lecture, December 11, 1951.) Copyright The Nobel Foundation, 1952. Elsevier Publishing Company, Amsterdam. From *The Development of High-Energy Accelerators,* edited by M. Stanley Livingston, Classics of Science, Volume 3, Dover Publications Inc., New York, 1966.
8. *Particle Accelerators* by Robert R. Wilson from *Scientific American*, March 1958. Reprinted with permission. Copyright © 1958 by Scientific American, Inc. All rights reserved.
9. *The Cyclotron as Seen by . . .* Cartoons by David L. Judd and Ronald G. MacKenzie, prepared for the International Conference on Isochronous Cyclotrons, Gatlinburg, Tenn., May, 1966. Reprinted from the proceedings of the conference (IEEE *Transactions in Nuclear Science,* vol.

NS-13, No. 4, August 1966) with the permission of the IEEE.

10. *CERN* by Jeremy Bernstein from *A Comprehensible World:* Essays on Science, copyright © 1964 by Jeremy Bernstein. Reprinted with permission of Random House, Inc. This article originally appeared in *The New Yorker.*
11. *New World of Nuclear Power* from *Introduction To Natural Science, Part I: The Physical Sciences* by V. L. Parsegian, pages 633–641, copyright © 1968 by Academic Press. Reprinted with permission.
12. *The Atomic Nucleus* by R. E. Peierls from *Scientific American,* January 1959. Reprinted with permission. Copyright © 1959 by Scientific American, Inc. All rights reserved.
13. *Power from the Stars* by Ralph E. Lapp from *Roads to Discovery,* pages 159–170, copyright © 1960 by Ralph E. Lapp. Harper & Row, Publishers, New York.
14. *Success* by Laura Fermi from *Atoms In The Family,* copyright 1954 by the University of Chicago Press, pages 190–199. Reprinted with permission.
15. *The Nuclear Energy Revolution—1966,* by Alvin M. Weinberg and Gale Young. *Proceedings of the National Academy of Science,* Vol. 57, No. 1, pp. 1–15, January 1967. Research sponsored by the U.S. Atomic Energy Commission under contract with the Union Carbide Corporation. Reprinted with permission.
16. *Conservation Laws* by Kenneth W. Ford from *The World of Elementary Particles,* copyright © 1963 by Blaisdell Publishing Company, a division of Ginn and Company, Watham, Massachusetts, pages 81–112. Reprinted with permission.
17. *The Fall of Parity* by Martin Gardner from *The Ambidextrous Universe,* copyright © 1964 by Martin Gardner. Reprinted with permission of Basic Books, Inc., New York and Penguin Books Ltd.
18. *Can Time Go Backward?* by Martin Gardner from *Scientific American,* January 1967, copyright © 1967 by Scientific American, Inc. Reprinted with permission. All rights reserved. Available separately at 25¢ each as Offprint No. 309 from W. H. Freeman and Company, 660 Market Street, San Francisco, California 94104.
19. *A Report to the Secretary of War,* by J. Franck, D. J. Hughes, J. J. Nickson, E. Rabinowitch, G. T. Seaborg, J. C. Stearns, L. Szilard, June 1945, Chapter 3 of *The Atomic Age,* edited by Morton Grodzins and Eugene Rabinowitch, copyright © 1963 by Basic Books, Inc., New York. Reprinted with permission.
20. *The Privilege of Being a Physicist* by Victor F. Weisskopf from *Physics Today,* August 1969, pages 39–43, copyright © 1969. Reprinted with permission.
21. *Calling All Stars* by Leo Szilard from *Voice of the Dolphins,* pages 105–111, copyright © 1961 by Leo Szilard. Reprinted by permission of Simon and Schuster, New York.
22. *Tasks for a World Without War* by Harrison Brown from *Daedalus,* Fall 1960, Journal of the American Academy of Arts and Sciences, Boston, pages 1029–1038. Reprinted with permission.
23. *One Scientist and His View of Science* by Leopold Infeld from *Quest,* copyright 1941 by Leopold Infeld. Reprinted by permission of Russell & Volkening, New York.
24. *Development of the Space-Time View of Quantum Electrodynamics* by Richard P. Feynman (Nobel Lecture, December 11, 1965), copyright © The Nobel Foundation 1966, Elsevier Publishing Company, Amsterdam. Reprinted with permission from *Science,* August 12, 1966, Volume 153, Number 3737, pages 609–708.
25. *The Relation of Mathematics to Physics,* by Richard P. Feynman from *The Character of Physical Law,* pages 55–57, British Broadcasting Corporation, London, copyright © 1965 by Richard P. Feynman. Reprinted with permission of the author and the M.I.T. Press, Cambridge, Massachusetts.
26. *Where Do We Go From Here* by Arthur E. Ruark from *Physics Today,* September 1969, pages 25–28, copyright 1969. Reprinted with permission.



This is not a physics textbook. Rather, it is a physics reader, a collection of some of the best articles and book passages on physics. A few are on historic events in science, others contain some particularly memorable description of what physicists do; still others deal with philosophy of science, or with the impact of scientific thought on the imagination of the artist.

There are old and new classics, and also some little-known publications; many have been suggested for inclusion because some teacher or physicist remembered an article with particular fondness. The majority of articles is not drawn from scientific papers of historic importance themselves, because material from many of these is readily available, either as quotations in the Project Physics text or in special collections.

This collection is meant for your browsing. If you follow your own reading interests, chances are good that you will find here many pages that convey the joy these authors have in their work and the excitement of their ideas. If you want to follow up on interesting excerpts, the source list at the end of the reader will guide you for further reading.

Reader 6

Table of Contents

1 **Rutherford** 1
Charles P. Snow

2 **The Nature of the Alpha Particle** 19
Ernest Rutherford and T. Royds

3 **Some Personal Notes on the Search for the Neutron** 25
Sir James Chadwick

4 **Antiprotons** 32
Owen Chamberlain, Emilio Segré, Clyde E. Wiegand, and Thomas Ypsilantis

5 **The Tracks of Nuclear Particles** 35
Herman Yagoda

6 **The Spark Chamber** 43
Gerard K. O'Neill

7 **The Evolution of the Cyclotron** 51
Ernest O. Lawrence

8 **Particle Accelerators** 65
Robert K. Wilson

9 **The Cyclotron As Seen By . . .** 77
David C. Judd and Ronald MacKenzie

10 **CERN** 83
Jeremy Bernstein

11 **The World of New Atoms and of Ionizing Radiations** 95
V. Lawrence Parsegian, Alan S. Meltzer, Abraham S. Luchins, K. Scott Kinerson

12 **The Atomic Nucleus** 103
Rudolf E. Peierls

13 **Power from the Stars** 109
Ralph E. Lapp

14 **Success** 123
Laura Fermi

15 **The Nuclear Energy Revolution** 135
Alvin M. Weinberg and Gale Young

16 **Conservation Laws** 141
Kenneth W. Ford

17 **The Fall of Parity** 175
Martin Gardner

18 **Can Time Go Backward?** 193
Martin Gardner

19 **A Report to the Secretary of War** 201
James Franck, Donald J. Hughes, J. I. Nickson, Eugene Rabinowitch, Glenn T. Seaborg, Joyce C. Stearns, Leo Szilard

20 **The Privilege of Being a Physicist** 211
Victor F. Weisskopf

21 **Calling All Stars** 221
Leo Szilard

22 **Tasks for a World Without War** 227
Harrison Brown

23 **One Scientist and His View of Science** 237
Leopold Infeld

24 **The Development of the Space-Time View of Quantum Electrodynamics** 241
Richard P. Feynman

25 **The Relation of Mathematics to Physics** 251
Richard P. Feynman

26 **Where Do We Go From Here?** 253
Arthur E. Ruark

C. P. Snow's highly personal account of Ernest Rutherford is based partly on Snow's research work in the Cavendish Laboratory while Rutherford was director.

1 Rutherford

Charles P. Snow

Chapter from his book, *Variety of Men*, published in 1967.

In 1923, at the meeting of the British Association for the Advancement of Science in Liverpool, Rutherford announced, at the top of his enormous voice: "We are living in the heroic age of physics." He went on saying the same thing, loudly and exuberantly, until he died, fourteen years later.

The curious thing was, all he said was absolutely true. There had never been such a time. The year 1932 was the most spectacular year in the history of science. Living in Cambridge, one could not help picking up the human, as well as the intellectual, excitement in the air. James Chadwick, grey-faced after a fortnight of work with three hours' sleep a night, telling the Kapitsa Club (to which any young man was so proud to belong) how he had discovered the neutron; P. M. S. Blackett, the most handsome of men, not quite so authoritative as usual, because it seemed too good to be true, showing plates which demonstrated the existence of the positive electron; John Cockcroft, normally about as much given to emotional

display as the Duke of Wellington, skimming down King's Parade and saying to anyone whose face he recognized: "We've split the atom! We've split the atom!"

It meant an intellectual climate different in kind from anything else in England at the time. The tone of science was the tone of Rutherford: magniloquently boastful—boastful because the major discoveries were being made—creatively confident, generous, argumentative, lavish, and full of hope. The tone differed from the tone of literary England as much as Rutherford's personality differed from that of T. S. Eliot. During the twenties and thirties Cambridge was the metropolis of experimental physics for the entire world. Even in the late nineteenth century, during the professorships of Clerk Maxwell and J. J. Thomson, it had never quite been that. "You're always at the crest of the wave," someone said to Rutherford. "Well, after all, I made the wave, didn't I?" Rutherford replied.

I remember seeing him a good many times before I first spoke to him. I was working on the periphery of physics at the time, and so didn't come directly under him. I already knew that I wanted to write novels, and that was how I should finish, and this gave me a kind of ambivalent attitude to the scientific world; but, even so, I could not avoid feeling some sort of excitement, or enhancement of interest, whenever I saw Rutherford walking down Free School Lane.

He was a big, rather clumsy man, with a substantial bay-window that started in the middle of the chest. I should guess that he was less muscular than at first sight he looked. He had large staring blue eyes and a damp and

pendulous lower lip. He didn't look in the least like an intellectual. Creative people of his abundant kind never do, of course, but all the talk of Rutherford looking like a farmer was unperceptive nonsense. His was really the kind of face and physique that often goes with great weight of character and gifts. It could easily have been the soma of a great writer. As he talked to his companions in the street, his voice was three times as loud as any of theirs, and his accent was bizarre. In fact, he came from the very poor: his father was an odd-job man in New Zealand and the son of a Scottish emigrant. But there was nothing Antipodean or Scottish about Rutherford's accent; it sounded more like a mixture of West Country and Cockney.

In my first actual meeting with him, perhaps I could be excused for not observing with precision. It was early in 1930; I had not yet been elected a Fellow of my own college, and so had put in for the Stokes studentship at Pembroke. One Saturday afternoon I was summoned to an interview. When I arrived at Pembroke, I found that the short list contained only two, Philip Dee and me. Dee was called in first; as he was being interviewed, I was reflecting without pleasure that he was one of the brightest of Rutherford's bright young men.

Then came my turn. As I went in, the first person I saw, sitting on the right hand of the Master, was Rutherford himself. While the Master was taking me through my career, Rutherford drew at his pipe, not displaying any excessive interest in the proceedings. The Master came to the end of his questions, and said: "Professor Rutherford?"

Rutherford took out his pipe and turned on to me an eye which was blue, cold and bored. He was the most spontaneous of men; when he felt bored he showed it. That afternoon he felt distinctly bored. Wasn't his man, and a very good man, in for this job? What was this other fellow doing there? Why were we all wasting our time?

He asked me one or two indifferent questions in an irritated, impatient voice. What was my present piece of work? What could spectroscopy tell us anyway? Wasn't it just "putting things into boxes?"

I thought that was a bit rough. Perhaps I realized that I had nothing to lose. Anyway, as cheerfully as I could manage, I asked if he couldn't put up with a few of us not doing nuclear physics. I went on, putting a case for my kind of subject.

A note was brought round to my lodgings that evening. Dee had got the job. The electors wished to say that either candidate could properly have been elected. That sounded like a touch of Cambridge politeness, and I felt depressed. I cheered up a day or two later when I heard that Rutherford was trumpeting that I was a young man of spirit. Within a few months he backed me for another studentship. Incidentally, Dee was a far better scientist than I was or could have been, and neither Rutherford nor anyone else had been unjust.

From that time until he died, I had some opportunities of watching Rutherford at close quarters. Several of my friends knew him intimately, which I never did. It is a great pity that Tizard or Kapitsa, both acute psychological observers, did not write about him at length. But I belonged to a dining club which he attended, and I think I

had serious conversations with him three times, the two of us alone together.

The difficulty is to separate the inner man from the Rutherfordiana, much of which is quite genuine. From behind a screen in a Cambridge tailor's, a friend and I heard a reverberating voice: "That shirt's too tight round the neck. Every day I grow in girth. *And* in mentality." Yet his physical make-up was more nervous than it seemed. In the same way, his temperament, which seemed exuberantly powerful, massively simple, rejoicing with childish satisfaction in creation and fame, was not quite so simple as all that. His was a personality of Johnsonian scale. As with Johnson, the façade was overbearing and unbroken. But there were fissures within.

No one could have enjoyed himself more, either in creative work or the honors it brought him. He worked hard, but with immense gusto; he got pleasure not only from the high moments, but also from the hours of what to others would be drudgery, sitting in the dark counting the alpha particle scintillations on the screen. His insight was direct, his intuition, with one curious exception, infallible. No scientist has made fewer mistakes. In the corpus of his published work, one of the largest in scientific history, there was nothing he had to correct afterwards. By thirty he had already set going the science of nuclear physics—single-handed, as a professor on five hundred pounds a year, in the isolation of late-Victorian Montreal. By forty, now in Manchester, he had found the structure of the atom—on which all modern nuclear physics depends.

It was an astonishing career, creatively active until the month he died. He was born very poor, as I have said.

New Zealand was, in the 1880's, the most remote of provinces, but he managed to get a good education; enough of the old Scottish tradition had percolated there, and he won all the prizes. He was as original as Einstein, but unlike Einstein he did not revolt against formal instruction; he was top in classics as well as in everything else. He started research—on the subject of wireless waves—with equipment such as one might rustle up today in an African laboratory. That did not deter him: "I could do research at the North Pole," he once proclaimed, and it was true. Then he was awarded one of the 1851 overseas scholarships (which later brought to England Florey, Oliphant, Philip Bowden, a whole series of gifted Antipodeans). In fact, he got the scholarship only because another man, placed above him, chose to get married: with the curious humility that was interwoven with his boastfulness, he was grateful all of his life. There was a proposal, when he was Lord Rutherford, President of the Royal Society, the greatest of living experimental scientists, to cut down these scholarships. Rutherford was on the committee. He was too upset to speak: at last he blurted out:

"If it had not been for them, I shouldn't have been." That was nonsense. Nothing could have stopped him. He brought his wireless work to Cambridge, anticipated Marconi, and then dropped it because he saw a field—radioactivity—more scientifically interesting.

If he had pushed on with wireless, incidentally, he couldn't have avoided becoming rich. But for that he never had time to spare. He provided for his wife and daughter, they lived in comfortable middle-class houses,

and that was all. His work led directly to the atomic energy industry spending, within ten years of his death, thousands of millions of pounds. He himself never earned, or wanted to earn, more than a professor's salary—about £1,600 a year at the Cavendish in the thirties. In his will he left precisely the value of his Nobel Prize, then worth £7,000. Of the people I am writing about, he died much the poorest *: even G. H. Hardy, who by Rutherford's side looked so ascetic and unworldly, happened not to be above taking an interest in his investments.

As soon as Rutherford got on to radioactivity, he was set on his life's work. His ideas were simple, rugged, material: he kept them so. He thought of atoms as though they were tennis balls. He discovered particles smaller than atoms, and discovered how they moved or bounced. Sometimes the particles bounced the wrong way. Then he inspected the facts and made a new but always simple picture. In that way he moved, as certainly as a sleepwalker, from unstable radioactive atoms to the discovery of the nucleus and the structure of the atom.

In 1919 he made one of the significant discoveries of all time: he broke up a nucleus of nitrogen by a direct hit from an alpha particle. That is, man could get inside the atomic nucleus and play with it if he could find the right projectiles. These projectiles could either be provided by radioactive atoms or by ordinary atoms speeded up by electrical machines.

The rest of that story leads to the technical and military history of our time. Rutherford himself never built the great machines which have dominated modern parti-

* One has to leave Stalin out of this comparison.

cle physics, though some of his pupils, notably Cockcroft, started them. Rutherford himself worked with bizarrely simple apparatus: but in fact he carried the use of such apparatus as far as it would go. His researches remain the last supreme single-handed achievement in fundamental physics. No one else can ever work there again—in the old Cavendish phrase—with sealing wax and string.

It was not done without noise: it was done with anger and storms—but also with an overflow of creative energy, with abundance and generosity, as though research were the easiest and most natural avocation in the world. He had deep sympathy with the creative arts, particularly literature; he read more novels than most literary people manage to do. He had no use for critics of any kind. He felt both suspicion and dislike of the people who invested scientific research or any other branch of creation with an aura of difficulty, who used long, methodological words to explain things which he did perfectly by instinct. "Those fellows," he used to call them. "Those fellows" were the logicians, the critics, the metaphysicians. They were clever; they were usually more lucid than he was; in argument against them he often felt at a disadvantage. Yet somehow they never produced a serious piece of work, whereas he was the greatest experimental scientist of the age.

I have heard larger claims made for him. I remember one discussion in particular, a year or two after his death, by half-a-dozen men, all of whom had international reputations in science. Darwin was there: G. I. Taylor: Fowler and some others. Was Rutherford the greatest experimental scientist since Michael Faraday? Without any doubt.

Greater than Faraday? Possibly so. And then—it is interesting, as it shows the anonymous Tolstoyan nature of organized science—how many years' difference would it have made if he had never lived? How much longer before the nucleus would have been understood as we now understand it? Perhaps ten years. More likely only five.

Rutherford's intellect was so strong that he would, in the long run, have accepted that judgment. But he would not have liked it. His estimate of his own powers was realistic, but if it erred at all, it did not err on the modest side. "There is no room for this particle in the atom as designed by *me*," I once heard him assure a large audience. It was part of his nature that, stupendous as his work was, he should consider it 10 per cent more so. It was also part of his nature that, quite without acting, he should behave constantly as though he were 10 per cent larger than life. Worldly success? He loved every minute of it: flattery, titles, the company of the high official world. He said in a speech: "As I was standing in the drawing-room at Trinity, a *clergyman* came in. And I said to him: 'I'm Lord Rutherford.' And he said to me: 'I'm the Archbishop of York.' And I don't suppose either of us believed the other."

He was a great man, a very great man, by any standards which we can apply. He was not subtle: but he was clever as well as creatively gifted, magnanimous (within the human limits) as well as hearty. He was also superbly and magnificently vain as well as wise—the combination is commoner than we think when we are young. He enjoyed a life of miraculous success. On the whole he enjoyed his own personality. But I am sure that, even quite late in his life, he felt stabs of a sickening insecurity.

Somewhere at the roots of that abundant and creative nature there was a painful, shrinking nerve. One has only to read his letters as a young man to discern it. There are passages of self-doubt which are not to be explained completely by a humble colonial childhood and youth. He was uncertain in secret, abnormally so for a young man of his gifts. He kept the secret as his personality flowered and hid it. But there was a mysterious diffidence behind it all. He hated the faintest suspicion of being patronized, even when he was a world figure. Archbishop Lang was once tactless enough to suggest that he supposed a famous scientist had no time for reading. Rutherford immediately felt that he was being regarded as an ignorant roughneck. He produced a formidable list of his last month's reading. Then, half innocently, half malevolently: "And what do you manage to read, your Grice?" "I am afraid," said the Archbishop, somewhat out of his depth, "that a man in my position really doesn't have the leisure. . . ." "Ah, yes, your Grice," said Rutherford in triumph, "it must be a dog's life! It must be a dog's life!"

Once I had an opportunity of seeing that diffidence face to face. In the autumn of 1934 I published my first novel, which was called *The Search* and the background of which was the scientific world. Not long after it came out, Rutherford met me in King's Parade. "What have you been doing to us, young man?" he asked vociferously. I began to describe the novel, but it was not necessary; he announced that he had read it with care. He went on to invite, or rather command, me to take a stroll with him round the Backs. Like most of my scientific friends, he

was good-natured about the book, which has some descriptions of the scientific experience which are probably somewhere near the truth. He praised it. I was gratified. It was a sunny October afternoon. Suddenly he said: "I didn't like the erotic bits. I suppose it's because we belong to different generations."

The book, I thought, was reticent enough. I did not know how to reply.

In complete seriousness and simplicity, he made another suggestion. He hoped that I was not going to write all my novels about scientists. I assured him that I was not—certainly not another for a long time.

He nodded. He was looking gentler than usual, and thoughtful. "It's a small world, you know," he said. He meant the world of science. "Keep off us as much as you can. People are bound to think that you are getting at some of us. And I suppose we've all got things that we don't want anyone to see."

I mentioned that his intuitive foresight went wrong just once. As a rule, he was dead right about the practical applications of science, just as much as about the nucleus. But his single boss shot sounds ironic now. In 1933 he said, in another address to the British Association, "These transformations of the atom are of extraordinary interest to scientists, but we cannot control atomic energy to an extent which would be of any value commercially, and I believe we are not likely ever to be able to do so. A lot of nonsense has been talked about transmutations. Our interest in the matter is purely scientific."

That statement, which was made only nine years be-

fore the first pile worked, was not intended to be either optimistic or pessimistic. It was just a forecast, and it was wrong.

That judgment apart, people outside the scientific world often felt that Rutherford and his kind were optimistic—optimistic right against the current of the twentieth century literary-intellectual mood, offensively and brazenly optimistic. This feeling was not quite unjustified, but the difference between the scientists and the non-scientists was more complex than that. When the scientists talked of the individual human condition, they did not find it any more hopeful than the rest of us. Does anyone really imagine that Bertrand Russell, G. H. Hardy, Rutherford, Blackett and the rest were bemused by cheerfulness as they faced their own individual state? Very few of them had any of the consolations of religion: they believed, with the same certainty that they believed in Rutherford's atom, that they were going, after this mortal life, into annihilation. Several of them were men of deep introspective insight. They did not need teaching anything at all about the existential absurdity.

Nevertheless it is true that, of the kinds of people I have lived among, the scientists were much the happiest. Somehow scientists were buoyant at a time when other intellectuals could not keep away despair. The reasons for this are not simple. Partly, the nature of scientific activity, its complete success on its own terms, is itself a source of happiness; partly, people who are drawn to scientific activity tend to be happier in temperament than other clever men. By the nature of their vocation and also by the nature of their own temperament, the scientists did

not think constantly of the individual human predicament. Since they could not alter it, they let it alone. When they thought about people, they thought most of what could be altered, not what couldn't. So they gave their minds not to the individual condition but to the social one.

There, science itself was the greatest single force for change. The scientists were themselves part of the deepest revolution in human affairs since the discovery of agriculture. They could accept what was happening, while other intellectuals shrank away. They not only accepted it, they rejoiced in it. It was difficult to find a scientist who did not believe that the scientific-technical-industrial revolution, accelerating under his eyes, was not doing incomparably more good than harm.

This was the characteristic optimism of scientists in the twenties and thirties. Is it still? In part, I think so. But there has been a change.

In the Hitler war, physicists became the most essential of military resources: radar, which occupied thousands of physicists on both sides, altered the shape of the war, and the nuclear bomb finished large scale "conventional" war for ever. To an extent, it had been foreseen by the mid-thirties that if it came to war (which a good many of us expected) physicists would be called on from the start. Tizard was a close friend of Rutherford's, and kept him informed about the prospects of RDF (as radar was then called). By 1938 a number of the Cavendish physicists had been secretly indoctrinated. But no one, no one at all, had a glimmering of how, for a generation afterwards, a high percentage of all physicists in the

United States, the Soviet Union, this country, would remain soldiers-not-in-uniform. Mark Oliphant said sadly, when the first atomic bomb was dropped: "This has killed a beautiful subject." Intellectually that has turned out not to be true: but morally there is something in it. Secrecy, national demands, military influence, have sapped the moral nerve of physics. It will be a long time before the climate of Cambridge, Copenhagen, Göttingen in the twenties is restored: or before any single physicist can speak to all men with the calm authority of Einstein or Bohr. That kind of leadership has now passed to the biologists, who have so far not been so essential to governments. It will be they, I think, who are likely to throw up the great scientific spokesmen of the next decades. If someone now repeated Gorki's famous question, "Masters of culture, which side are you on?" it would probably be a biologist who spoke out for his fellow human beings.

In Rutherford's scientific world, the difficult choices had not yet formed themselves. The liberal decencies were taken for granted. It was a society singularly free from class or national or racial prejudice. Rutherford called himself alternatively conservative or non-political, but the men he wanted to have jobs were those who could do physics. Niels Bohr, Otto Hahn, Georg von Hevesy, Hans Geiger were men and brothers, whether they were Jews, Germans, Hungarians—men and brothers whom he would much rather have near him than the Archbishop of Canterbury or one of "those fellows" or any damned English philosopher. It was Rutherford who, after 1933, took the lead in opening English academic life to Jewish refugees. In fact, scientific society was wide open, as it may

not be again for many years. There was coming and going among laboratories all over the world, including Russia. Peter Kapitsa, Rutherford's favorite pupil, contrived to be in good grace with the Soviet authorities and at the same time a star of the Cavendish.

He had a touch of genius: in those days, before life sobered him, he had also a touch of the inspired Russian clown. He loved his own country, but he distinctly enjoyed backing both horses, working in Cambridge and taking his holidays in the Caucasus. He once asked a friend of mine if a foreigner could become an English peer; we strongly suspected that his ideal career would see him established simultaneously in the Soviet Academy of Sciences and as Rutherford's successor in the House of Lords.

At that time Kapitsa attracted a good deal of envy, partly because he could do anything with Rutherford. He called Rutherford "the Crocodile," explaining the crocodile means "father" in Russian, which it doesn't, quite: he had Eric Gill carve a crocodile on his new laboratory. He flattered Rutherford outrageously, and Rutherford loved it. Kapitsa could be as impertinent as a Dostoevskian comedian: but he had great daring and scientific insight. He established the club named after him (which again inspired some envy) : it met every Tuesday night, in Kapitsa's rooms in Trinity, and was deliberately kept small, about thirty, apparently because Kapitsa wanted to irritate people doing physical subjects he disapproved of. We used to drink large cups of milky coffee immediately after hall (living was fairly simple, and surprisingly non-alcoholic, in scientific Cambridge), and someone gave a talk—often a dramatic

one, like Chadwick's on the neutron. Several of the major discoveries of the thirties were first heard in confidence in that room. I don't think that the confidence was ever broken.

I myself enjoyed the one tiny scientific triumph of my life there. At the time Kapitsa barely tolerated me, since I did spectroscopy, a subject he thought fit only for bank clerks: in fact I had never discovered why he let me join. One night I offered to give a paper outside my own subject, on nuclear spin, in which I had been getting interested: I didn't know much about it, but I reckoned that most of the Cavendish knew less. The offer was unenthusiastically accepted. I duly gave the paper. Kapitsa looked at me with his large blue eyes, with a somewhat unflattering astonishment, as at a person of low intelligence who had contrived inadvertently to say something interesting. He turned to Chadwick, and said incredulously, "Jimmy, I believe there *is* something in this."

It was a personal loss to Rutherford when Kapitsa, on one of his holiday trips to Russia, was told by the Soviet bosses, politely but unyieldingly, that he must stay: he was too valuable, they wanted his services full-time. After a while Kapitsa made the best of it. He had always been a patriotic Russian: though both he and his wife came from the upper middle-class, if there was such a class in old Russia (his father was a general in the Tsarist engineering corps), he took a friendly attitude to the revolution. All that remained steady, though I don't think he would mind my saying that his enthusiasm for Stalin was not unqualified. Still, Kapitsa threw all his gifts into his new work in the cause of Soviet science. It was only then that we, who

had known him in Cambridge, realized how strong a character he was: how brave he was: and fundamentally what a good man. His friendship with Cockcroft and others meant that the link between Soviet and English science was never quite broken, even in the worst days. Only great scientists like Lev Landau can say in full what he has done for science in his own country. If he hadn't existed, the world would have been worse: that is an epitaph that most of us would like and don't deserve.

Between Leningrad and Cambridge, Kapitsa oscillated. Between Copenhagen and Cambridge there was a stream of travellers, all the nuclear physicists of the world. Copenhagen had become the second scientific metropolis on account of the personal influence of one man, Niels Bohr, who was complementary to Rutherford as a person—patient, reflective, any thought hedged with Proustian qualifications—just as the theoretical quantum physics of which he was the master was complementary to Rutherford's experimental physics. He had been a pupil of Rutherford's, and they loved and esteemed each other like father and son. (Rutherford was a paterfamilias born, and the death of his only daughter seems to have been the greatest sorrow of his personal life. In his relations with Bohr and Kapitsa and others, there was a strong vein of paternal emotion diverted from the son he never had.) But, strong as Rutherford's liking for Bohr was, it was not strong enough to put up with Bohr's idea of a suitable length for a lecture. In the Cavendish lecture room, Bohr went past the hour; Rutherford began to stir. Bohr went past the hour and a half; Rutherford began plucking at his sleeve and muttering in a stage whisper about "another

five minutes." Blandly, patiently, determined not to leave a qualification unsaid, as indefatigable as Henry James in his last period, Bohr went past the two hours; Rutherford was beginning to trumpet about "bringing the lecture to a close." Soon they were both on their feet at once.

Rutherford died suddenly when he was age sixty-six, still in full vigor. He died not only suddenly, but of something like a medical accident: he had a strangulated hernia. There was no discernible reason why he should not have lived into old age.

It was a sunny, tranquil October morning, the kind of day on which Cambridge looks so beautiful. I had just arrived at the crystallographic laboratory, one of the buildings in the old Cavendish muddle; why I was there I don't remember, nor whom I was talking to, except that it happened not to be Bernal. Someone put his head round the door and said: "The Professor's dead."

I don't think anyone said much more. We were stupefied rather than miserable. It did not seem in the nature of things.

Rutherford reports on his ingenious experiments proving that the alpha particle is a charged helium atom.

2 The Nature of the Alpha Particle

Ernest Rutherford and T. Royds

A paper in *Philosophical Magazine*, published in 1909.

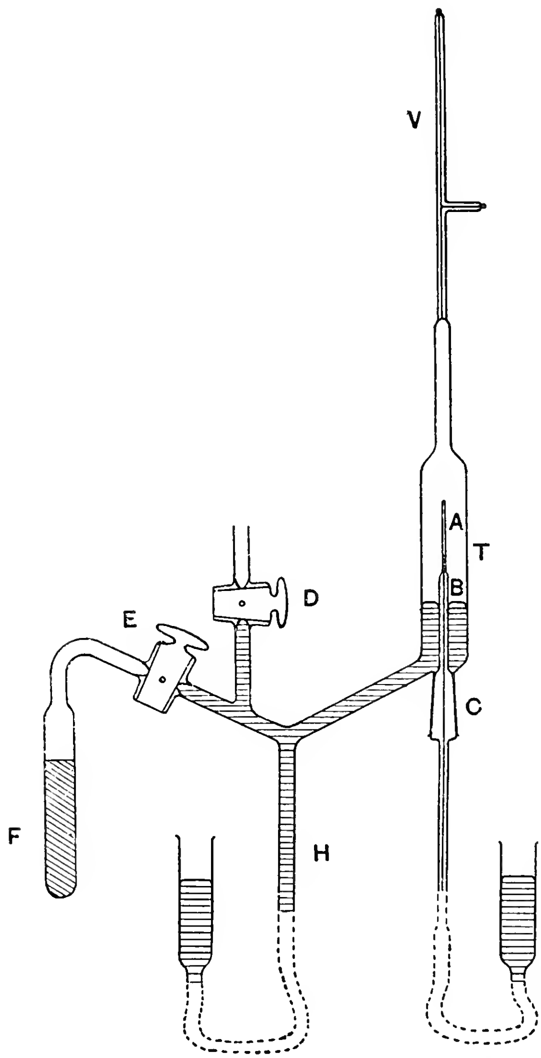
THE experimental evidence collected during the last few years has strongly supported the view that the α particle is a charged helium atom, but it has been found exceedingly difficult to give a decisive proof of the relation. In recent papers, Rutherford and Geiger † have supplied still further evidence of the correctness of this point of view. The number of α particles from one gram of radium have been counted, and the charge carried by each determined. The values of several radioactive quantities, calculated on the assumption that the α particle is a helium atom carrying two unit charges, have been shown to be in good agreement with the experimental numbers. In particular, the good agreement between the calculated rate of production of helium by radium and the rate experimentally determined by Sir James Dewar ‡, is strong evidence in favour of the identity of the α particle with the helium atom.

The methods of attack on this problem have been largely indirect, involving considerations of the charge carried by the helium atom and the value of e/m of the α particle. The proof of the identity of the α particle with the helium atom is incomplete until it can be shown that the α particles, accumulated quite independently of the matter from which they are expelled, consist of helium. For example, it might be argued that the appearance of helium in the radium emanation was a result of the expulsion of the α particle, in the same way that the appearance of radium A is a consequence of the expulsion of an α particle from the emanation. If one atom of helium appeared for each α particle expelled, calculation and experiment might still agree, and yet the α particle itself might be an atom of hydrogen or of some other substance.

We have recently made experiments to test whether helium appears in a vessel into which the α particles have been fired, the active matter itself being enclosed in a vessel sufficiently thin to allow the α particles to escape, but impervious to the passage of helium or other radioactive products.

* Communicated by the Authors.
† Proc. Roy. Soc. A. lxxxi. pp. 141–173 (1908).
‡ Proc. Roy. Soc. A. lxxxi. p. 280 (1908).

The experimental arrangement is clearly seen in the figure The equilibrium quantity of emanation from about 140 milligrams of radium was purified and compressed by means of a



mercury-column into a fine glass tube A about 1·5 cms. long. This fine tube, which was sealed on a larger capillary tube B, was sufficiently thin to allow the α particles from the emanation and its products to escape, but sufficiently strong to

withstand atmospheric pressure. After some trials, Mr. Baumbach succeeded in blowing such fine tubes very uniform in thickness. The thickness of the wall of the tube employed in most of the experiments was less than $\frac{1}{100}$ mm., and was equivalent in stopping power of the α particle to about 2 cms. of air. Since the ranges of the α particles from the emanation and its products radium A and radium C are 4·3, 4·8, and 7 cms. respectively, it is seen that the great majority* of the α particles expelled by the active matter escape through the walls of the tube. The ranges of the α particles after passing through the glass were determined with the aid of a zinc-sulphide screen. Immediately after the introduction of the emanation the phosphorescence showed brilliantly when the screen was close to the tube, but practically disappeared at a distance of 3 cms. After an hour, bright phosphorescence was observable at a distance of 5 cms. Such a result is to be expected. The phosphorescence initially observed was due mainly to the α particles of the emanation and its product radium A (period 3 mins.). In the course of time the amount of radium C, initially zero, gradually increased, and the α radiations from it of range 7 cms. were able to cause phosphorescence at a greater distance.

The glass tube A was surrounded by a cylindrical glass tube T, 7·5 cms. long and 1·5 cms. diameter, by means of a ground-glass joint C. A small vacuum-tube V was attached to the upper end of T. The outer glass tube T was exhausted by a pump through the stopcock D, and the exhaustion completed with the aid of the charcoal tube F cooled by liquid air. By means of a mercury column H attached to a reservoir, mercury was forced into the tube T until it reached the bottom of the tube A.

Part of the α particles which escaped through the walls of the fine tube were stopped by the outer glass tube and part by the mercury surface. If the α particle is a helium atom, helium should gradually diffuse from the glass and mercury into the exhausted space, and its presence could then be detected spectroscopically by raising the mercury and compressing the gases into the vacuum-tube.

In order to avoid any possible contamination of the apparatus with helium, freshly distilled mercury and entirely new glass apparatus were used. Before introducing the emanation into A, the absence of helium was confirmed

* The α particles fired at a very oblique angle to the tube would be stopped in the glass. The fraction stopped in this way would be small under the experimental conditions.

experimentally. At intervals after the introduction of the emanation the mercury was raised, and the gases in the outer tube spectroscopically examined. After 24 hours no trace of the helium yellow line was seen ; after 2 days the helium yellow was faintly visible; after 4 days the helium yellow and green lines were bright; and after 6 days all the stronger lines of the helium spectrum were observed. The absence of the neon spectrum shows that the helium present was not due to a leakage of air into the apparatus.

There is, however, one possible source of error in this experiment. The helium may not be due to the α particles themselves, but may have *diffused* from the emanation through the thin walls of the glass tube. In order to test this point the emanation was completely pumped out of A, and after some hours a quantity of helium, about 10 times the previous volume of the emanation, was compressed into the same tube A.

The outer tube T and the vacuum-tube were removed and a fresh apparatus substituted. Observations to detect helium in the tube T were made at intervals, in the same way as before, but no trace of the helium spectrum was observed over a period of eight days.

The helium in the tube A was then pumped out and a fresh supply of emanation substituted. Results similar to the first experiment were observed. The helium yellow and green lines showed brightly after four days.

These experiments thus show conclusively that the helium could not have diffused through the glass walls, but must have been derived from the α particles which were fired through them. In other words, the experiments give a decisive proof that the α particle after losing its charge is an atom of helium.

Other Experiments.

We have seen that in the experiments above described helium was not observed in the outer tube in sufficient quantity to show the characteristic yellow line until two days had elapsed. Now the equilibrium amount of emanation from 100 milligrams of radium should produce helium at the rate of about ·03 c.mm. per day. The amount produced in one day, if present in the outer tube, should produce a bright spectrum of helium under the experimental conditions. It thus appeared probable that the helium fired into the glass must escape very slowly into the exhausted space, for if the helium escaped at once, the presence of helium should have

been detected a few hours after the introduction of the emanation.

In order to examine this point more closely the experiments were repeated, with the addition that a cylinder of thin sheet lead of sufficient thickness to stop the α particles was placed over the fine emanation tube. Preliminary experiments, in the manner described later, showed that the lead-foil did not initially contain a detectable amount of helium. Twenty-four hours after the introduction into the tube A of about the same amount of emanation as before, the yellow and green lines of helium showed brightly in the vacuum-tube, and after two days the whole helium spectrum was observed. The spectrum of helium in this case after one day was of about the same intensity as that after the fourth day in the experiments without the lead screen. It was thus clear that the lead-foil gave up the helium fired into it far more readily than the glass.

In order to form an idea of the rapidity of escape of the helium from the lead some further experiments were made. The outer cylinder T was removed and a small cylinder of lead-foil placed round the thin emanation-tube surrounded the air at atmospheric pressure. After exposure for a definite time to the emanation, the lead screen was removed and gested for helium as follows. The lead-foil was placed in a glass tube between two stopcocks. In order to avoid a possible release of the helium present in the lead by pumping out the air, the air was displaced by a current of pure electrolytic oxygen *. The stopcocks were closed and the tube attached to a subsidiary apparatus similar to that employed for testing for the presence of neon and helium in the gases produced by the action of the radium emanation on water (Phil. Mag. Nov. 1908). The oxygen was absorbed by charcoal and the tube then heated beyond the melting-point of lead to allow the helium to escape. The presence of helium was then spectroscopically looked for in the usual way. Using this method, it was found possible to detect the presence of helium in the lead which had been exposed for only four hours to the α rays from the emanation. After an exposure of 24 hours the helium yellow and green lines came out brightly. These experiments were repeated several times with similar results.

A number of blank experiments were made, using samples of the lead-foil which had not been exposed to the α rays, but in no case was any helium detected. In a similar way,

* That the air was completely displaced was shown by the absence of neon in the final spectrum.

the presence of helium was detected in a cylinder of tinfoil exposed for a few hours over the emanation-tube.

These experiments show that the helium does not escape at once from the lead, but there is on the average a period of retardation of several hours and possibly longer.

The detection of helium in the lead and tin foil, as well as in the glass, removes a possible objection that the helium might have been in some way present in the glass initially, and was liberated as a consequence of its bombardment by the α particles.

The use of such thin glass tubes containing emanation affords a simple and convenient method of examining the effect on substances of an intense α radiation quite independently of the radioactive material contained in the tube.

We can conclude with certainty from these experiments that the α particle after losing its charge is a helium atom. Other evidence indicates that the charge is twice the unit charge carried by the hydrogen atom set free in the electrolysis of water.

University of Manchester,
Nov. 13, 1908.

Chadwick reminisces on the period when he, as Rutherford's collaborator, searched for evidence of the neutron in the sealing-wax-and-string tradition of experimentation.

3 Some Personal Notes on the Search for the Neutron

Sir James Chadwick

Speech delivered before the 10th International Congress of History of Science at Cornell University, New York, in 1962.

Some personal notes on the search for the neutron.

A few months after his Bakerian Lecture of June 1920, in which he first mentioned what had been in his mind for some time — the possible existence of a neutral particle formed by the close combination of a proton and an electron, Rutherford invited me to join him in following up the experiments on the artificial disintegration of nitrogen which he had made in Manchester.

There were a number of reasons for this invitation, so welcome to me. Among them was that I had made some improvements in the technique of counting scintillations, better optical arrangements and a strict discipline; but also he wanted someone to talk to, to while away the tedium of working in darkness.

It was during the periods of waiting to begin counting that he expounded to me at length his views on the problems of nuclear structure, and in particular on the difficulty in seeing how complex nuclei could possibly build up if the only elementary particles available were the proton and the electron, and the need therefore to invoke the aid of the neutron. He freely admitted that much of this was pure speculation, and, being averse to speculation without some basis of experiment, he seldom mentioned these matters except in private discussion. Indeed, I believe that only on one occasion after the Bakerian Lecture did he again refer publicly to his views on the role of the neutron. He had not abandoned the idea, and he had completely converted me. From time to time in the course of the following years, sometimes together sometimes myself alone, we made experiments to find evidence of the neutron, both its formation and its emission from atomic nuclei. I shall mention some of the more respectable of these attempts; there were others which were so desperate, so far-fetched as to belong to the days of alchemy.

Immediately after the Bakerian Lecture Rutherford had asked J.L. Glasson to look for the production of neutrons when an electric discharge was passed through hydrogen, and a little later J.K. Roberts made a somewhat similar experiment. He did not really expect that any evidence of the neutron would turn up in this way, but it had to be tried. Both the mass of hydrogen and the voltages used in these experiments were quite trivial.

It seemed to me not too unreasonable to look at hydrogen in the normal state notwithstanding its apparent stability. If a close

combination of proton and electron were possible at all, it might take place spontaneously; and the neutron so formed might break up again under the action of the cosmic radiation. With Rutherford's approval, I tried, in 1923, to detect the emission of γ radiation from the formation of neutrons in a large mass of hydrogenous material, using an ionisation chamber and a point-counter as the means of detection.

A few years later, in 1928, Geiger and Müller devised what is now universally called the Geiger counter, which enormously increased the ability to detect γ radiation. Geiger very kindly sent me two of his new counters, as well as instructions for making them. Immediately, Rutherford and I used this new instrument to repeat the experiment with hydrogen. We went to all manner of tricks in the hope of finding some trace of the neutron. We also examined in the same way some of the rare gases, and any rare element we could lay our hands on, just in case any sign of the formation of the neutron or its emission might turn up. I mention these experiments only in a general way because some were quite wildly absurd.

After my first attempt in this way I had considered the possibility that the neutron could be formed, or exist, only in a strong electric field; and that perhaps one might find some evidence by firing fast protons into atoms, especially those of higher atomic number where some electrons were tightly bound. This was the vague idea behind the remark in a letter to Rutherford which is quoted in Eve's Life (p. 301) — "I think we shall have to make a real search for the neutron. I believe I have a scheme which may just work ...". I thought that at least 200.000 volts would be necessary

for the acceleration of the protons. No suitable transformer was available and, although Rutherford was mildly interested, there was no money to spend on such a wild scheme. [I might mention that the research grant of the Cavendish was about £2000 a year, little even in those days for the amount of work which had to be supported.] I persisted with the idea for a year or two, and in the intervals of other work I tried to find a way of applying Tesla voltages to the acceleration of ions in a discharge tube. I had quite inadequate facilities, and no experience in such matters. I wasted my time – but no money.

During our work on the disintegration of the lighter elements by α particles Rutherford and I had not been unmindful of the possibility of the emission of neutrons, especially from those elements which did not emit protons. We looked for ~~very~~ faint scintillations due to a radiation undeflected by a magnetic field. The only specific reference to the search for the neutron in this way was made in a paper published in 1929, some years after the experiments.

The case of beryllium was interesting for two reasons. It did not emit protons under α particle bombardment; and – though a false argument – the mineral beryl was known to contain an unusual amount of helium, suggesting that perhaps the Be nucleus split up under the action of the cosmic radiation into two α particles and a neutron.

This matter intrigued me off and on for some years. I bombarded beryllium with α particles, with β particles and with γ rays, generally using the scintillation method to detect any effect. In those days this was the only method of much use in the presence of the strong γ radiation of the radium active deposit, the chief source of radiation available to me. Quite early on, too early perhaps, I tried to devise suitable electrical methods of counting.

I failed. Later, when the valve amplifier method had been developed by Greinacher, and put into use in the Cavendish by Wynn Williams, I was also able to make a polonium source, small but just enough for the purpose. With Constable and Pollard, I had another look at beryllium, and for a short but exciting time we thought we had found some evidence of the neutron. But somehow the evidence faded away. I was still groping in the dark.

The first indication of the neutron came in the work of H.C. Webster on the γ radiation excited in beryllium by α particle collisions. I had had such work, the excitation of γ rays by bombarding light elements with α particles, in mind for some years. An attempt had been made by L.H. Bastings, but this failed, because the polonium source was too weak and the instrument of detection, the electroscope, too insensitive. When the Geiger counter became available Webster took up this quest, but his first efforts were not very rewarding — I was still short of polonium.

This deficiency was overcome by the kind intercession of Dr. Feather, then in Baltimore, and the generosity of Dr. C.F. Burnam and Dr. F. West of the Kelly Hospital. They sent me a number of old radon tubes which together contained what was, for me, a very large quantity of radium D and its product polonium. This gift was of immense value both immediately and later on.

In the meantime, Bothe and Becker had taken up this matter and they were the first to publish results. But Webster made a most interesting observation, that the radiation from beryllium which was emitted in the same direction as the incident α particles was more penetrating than the radiation emitted in the backward direction. This

fact, clearly established, excited me; it could only be readily explained if the radiation consisted of particles, and, from its penetrating power, of neutral particles. Believing that a neutral particle ~~should~~ would produce tracks, though very sparsely ionised, I suggested that he should pass the radiation into an expansion chamber. To our dismay, for we were convinced that a neutral particle of some kind was involved, no such tracks were to be seen. We were very puzzled; we did not know how to reconcile the observations.

This near-miss occurred in June 1931. Shortly afterwards Webster left Cambridge for Bristol. I decided to take up the matter afresh, but my preparations were delayed, perhaps fortunately, by a change of my working quarters to another part of the laboratory. Then one morning I read the communication of the Curie-Joliots in the Comptes Rendus, in which they reported a still more surprising property of the radiation from beryllium — a most startling property. Not many minutes afterwards Dr. Feather came to my room to tell me about this report, as astonished as I was. A little later that morning I told Rutherford. It was a custom of long standing that I should visit him about 11 a.m. to tell him any news of interest and to discuss the work in progress in the laboratory. As I told him about the Curie-Joliot observation and their views on it, I saw his growing ~~incredulity~~ amazement; and finally he burst out "I don't believe it". Such an impatient remark was utterly out of character, and in all my long association with him I recall no similar occasion. I mention it to emphasize the electrifying effect of the Curie-Joliot report. Of course, Rutherford agreed that one must believe the observations; the explanation was quite another matter.

It so happened that I was just ready

to begin experiment, for I had prepared a beautiful source of polonium from the Baltimore material. I started with an open mind, though naturally my thoughts were on the neutron. I was reasonably sure that the Curie-Joliot observations could not be ascribed to a kind of Compton effect, for I had looked for this more than once. I was convinced that there was something quite new as well as strange. A few days of strenuous work were sufficient to show that these strange effects were due to a neutral particle and to enable me to measure its mass — the neutron postulated by Rutherford in 1920 had at last revealed itself.

I trust that I shall not be misunderstood if I add a postscript to this story. It is unnecessary to record my satisfaction, and delight, that the long search for the neutron had, in the end, been successful. The decisive clue had indeed been supplied by others. This after all is not unusual; advances in knowledge are generally the result of the work of many minds and hands. But I could not help but feel that I ought to have arrived sooner. I could offer myself many excuses — lack of facilities, and so on. But beyond all excuses I had to admit, if only to myself, that I had failed to think deeply enough about the properties of the neutron, especially about those properties which would most clearly furnish evidence of its existence. It was a chastening thought. I consoled myself with the reflection that it is much more difficult to say the first word on any subject, however obvious it may later appear, than the last word — a commonplace reflection and perhaps only an excuse.

The authors establish the existence of antiprotons and explain their belief that there must be antineutrons.

4 Antiprotons

Owen Chamberlain, Emilio Segré, Clyde E. Wiegand, and Thomas J. Ypsilantis

From the periodical *Nature*, published in 1956.

SINCE the development of Dirac's theory of the electron and the brilliant confirmation of one of its most startling predictions by the discovery of the positron by Anderson, it has been assumed most likely that the proton would also have its charge conjugate, the antiproton. The properties that define the antiproton are: (1) charge equal to the electron charge (also in sign); (2) mass equal to the proton mass; (3) stability against spontaneous decay; (4) ability to become annihilated by interaction with a proton or neutron, probably generating pions and releasing in some manner the energy $2\ mc^2$; (5) generation in pairs with ordinary nucleons; (6) magnetic moment equal but opposite to that of the proton; (7) fermion of spin $\frac{1}{2}$. Not all these properties are independent, but all might ultimately be subjected to experiment.

In cosmic rays, where such antiprotons could appear, some events have been observed which could be due to antiprotons; but their interpretation is uncertain.

In order to generate antiprotons in the laboratory, an absolute lower limit of the necessary energy is $2\ mc^2 = 1{\cdot}88$ BeV.; but the mechanism of the collision and the conservation of momentum influence this lower limit, which becomes 5·6 BeV. if the process is a nucleon-nucleon collision, or 4·4 BeV. if the process is a two-step one with the formation of a pion in a nucleon–nucleon collision followed by a pion–nucleon collision in which the nucleon–antinucleon pair is generated. These thresholds can be lowered appreciably by internal motions of nucleons in the nucleus. (Energies are quoted in the laboratory system.)

When the Berkeley bevatron was planned, the goal of 6 BeV. was set, in the hope that this energy would be sufficient to create antiprotons.

The methods of detection of the antiproton can make use of any of the seven properties listed above. It seemed that (1), (2) and (3) might be the easiest to ascertain; (4) would also be highly desirable; whereas (5)–(7) are at present very difficult to observe.

There are classical methods of measuring charge and mass of a particle that go back in their origin to *J. J.* Thomson. They entail the simultaneous measurement on the same particle of any two of the quantities momentum, velocity or energy, which in turn can be obtained from the observation of electric or magnetic deflexions, time of flight, range, scattering in photographic emulsions, etc. As for the charge, it is sufficient to measure its sign and its absolute value in a rough way only, because it is assumed that it is an integral multiple of the electronic charge.

After a detailed discussion, it was decided that momentum p and velocity v constituted the most promising combination for ascertaining the mass. The first successful experiment[1] was performed at

the end of September 1955, as follows. The momentum was measured by passing the particles generated by bombardment of a copper target with 6·2 BeV. protons through two deflecting magnetic fields and two magnetic lenses. This ensemble let through only particles for which $p = 1{\cdot}19$ BeV./c, if their charge is equal to that of the electron, including sign. The velocity was measured by a time-of-flight measurement between two scintillation counters 40 ft. apart. The pulse-size in the scintillators showed that the particles were singly charged.

The chief difficulty of the experiment rests with the fact that the antiprotons are accompanied by many pions—44,000 pions per antiproton in the most favourable conditions. For this reason provision must be made for eliminating spurious background effects. One of the most important steps is the insertion in the beam of two Čerenkov counters: one that is activated by particles with $v/c = \beta > 0{\cdot}79$, and one of a special type that is activated by particles with $0{\cdot}75 < \beta < 0{\cdot}78$. Pions with $p = 1{\cdot}19$ BeV./c have $\beta = 0{\cdot}99$, while antiprotons of the same value of p have $\beta = 0{\cdot}78$, and their respective times of flight for an interval of 40 ft. are 40×10^{-9} sec. and 51×10^{-9} sec. Particles with β in the interval between 0·75 and 0·78 trigger the sweep of an oscilloscope in which the time of flight between two scintillation counters 40 ft. apart is displayed. This time of flight appears as the distance between the two 'pips' due to the traversal of the counters. From this time of flight the mass is determined with an accuracy of 10 per cent for each particle. Up to now, about 250 particles have been observed and the average mass is known to about 5 per cent. It is $1{,}840 \pm 90$ electron masses.

The functioning of the whole apparatus is checked by sending through it positive protons in a separate run. These are obtained from a subsidiary target, and their orbits are selected in such a way that they have the same momentum as the antiproton.

The particles are observable after a time of flight of 10^{-7} sec., which rules out particles with a mean life much shorter than 10^{-8} sec., in particular the known hyperons. These measurements are thus in agreement with points (1), (2) and (3) mentioned above, and the identification of the new particle with the antiproton is a natural one, although not absolutely established.

There are also some indications on the fourth point mentioned above, namely, the terminal process of the particle. Particles selected as antiprotons by the apparatus of ref. 1 were sent into a block of heavy glass and the Čerenkov radiation generated in it was measured[2]. This radiation does not correspond, of course, to the entirety of the energy released; actually it is only a small part of it. However, a calibration was performed, and from the pulse size the visible energy was estimated. Values up to 800 MeV. were found. This is consistent with the expected modes of annihilation for an antiproton, and with the energy it would throw into Čerenkov radiation in a detectable form; but it is not sufficient yet for positive identification on that score only.

Another type of observation on the terminal phenomenon accompanying the absorption of the antiproton was also performed[3] with the photographic plate technique. Particles of selected momentum obtained with an arrangement similar to that described in ref. 1 were slowed down by a copper absorber and finally stopped in a stack of photographic emulsions. Among a background of many pions one particle was found which has protonic mass, comes to rest and produces a star containing six black tracks, one grey proton, one pion of 58 MeV. and one minimum ionization track. The visible energy released is larger than 830 MeV. The total energy released cannot be known, because there are neutral particles emitted; but this amount of visible energy is also consistent with the annihilation of an antiproton.

Clearly many questions are raised by the new particle. Its identification should be further corroborated; it is important to study in detail its annihilation properties for complex nuclei and, possibly even more interesting, the annihilation with hydrogen and deuterium. In addition, the cross-section for nuclear interaction and the mechanism of production are clearly to be investigated.

The existence of the antiproton entails with virtual certainty the existence of the antineutron. Its experimental demonstration is a most interesting problem. Probably the neutron beam of the Berkeley bevatron contains an appreciable number of them, but their disentanglement from the ordinary neutrons appears a formidable task. It is likely that the best approach will be either: (1) to transform an antiproton into an antineutron by a collision with a proton; or (2) to convert an antineutron into an antiproton by collision with an ordinary neutron and detect either the final antineutron in (1) or the final antiproton in (2).

[1] Chamberlain, Segrè, Wiegand and Ypsilantis, *Phys. Rev.*, **100**, 947 (1955).

[2] Brabant, Cork, Horwitz, Moyer, Murray, Wallace and Wenzel, *Phys. Rev.* (in the press).

[3] Chamberlain, Chupp, Goldhaber, Segrè, Wiegand, and Amaldi, Baroni, Castagnoli, Franzinetti and Manfredini (to be published).

GIANT SHOWER OF MESONS is recorded in this photomicrograph of a small section of nuclear emulsion carried to a height of 106,000 feet by a Navy "Skyhook" balloon. At the top of the photomicrograph is the heavy track of an enormously energetic iron nucleus in the primary cosmic radiation. Above the nucleus is a "star" resulting from the collision of the iron nucleus and a nucleus in the emulsion. Below the star is a jet of about 40 pi mesons. To the left and right of the star are heavier fragments of the target nucleus.

Elementary particles can be studied by the traces they leave in photographic plates.

5 The Tracks of Nuclear Particles

Herman Yagoda

Article published in 1956 in the *Scientific American.*

A nuclear physicist studying the elementary particles of nature is in much the same position as an explorer trying to picture unknown animals from their tracks. The physicist never can see the particles themselves—only their footprints in a cloud chamber or a photographic plate. But from these tracks he deduces a particle's mass, movements, speed, lifetime and social impact on its fellows. By now the tracks of some members of the nuclear family are almost as familiar and readable as the footprints of a domestic animal. Interesting new tracks keep turning up, some strange, some predictable—the latest to make its appearance is that of the long-sought antiproton. It seems a timely moment to survey the scene and review the gallery of footprints that identify the members of the strange population in the nucleus of the atom.

We shall consider the tracks as they are recorded in photographic emulsions. It was in this medium that the existence of particles in the nucleus of the atom was first detected—through the fact that Henri Becquerel left some uranium near photographic film in a drawer. Becquerel noted simply that radioactive emanations from the uranium had fogged his film. That the "fog" might consist of a network of tracks was not discovered until 13 years later. In 1909 Otto Mügge of Germany exposed some film to tiny crystals of zircon, a feebly radioactive mineral. To study the faintly developed image he had to use a microscope, and he then noticed that there were fine linear tracks radiating from the crystals. Not long afterward the tracks of alpha particles emitted by radium were recorded in fine-grained emulsions at Lord Rutherford's famous laboratory in England.

When a charged particle travels through a photographic emulsion, it forms a latent image in the silver bromide grains, just as light does. In the case of the moving particle, the latent image results from ionization by the particle along its path. This image, marking the track of the particle, is then made visible by development of the emulsion in the usual way. So that fast particles may be brought to a stop within the emulsion, it is usually made as thick as possible. Emulsions used to track cosmic rays and high-energy particles from accelerators are often more than one millimeter thick—about 100 times thicker than in ordinary photographic film. The length of a particle's track in the emulsion must be measured precisely to determine its kinetic energy. Since the path slants into the emulsion, its length cannot be measured directly: it is computed by means of the Pythagorean theorem from the two measurable distances—the depth at which the particle comes to rest in the emulsion and the horizontal distance along the emulsion surface from the point of entry to the point directly above the end of the track.

At best the search for particle tracks in emulsions is slow, tedious work. It takes many hours or days of poring over the photographic plate with a microscope to find and trace the faint lines of silver grains. For this reason physicists long preferred to use cloud chambers for particle detection work. But the photographic plate has an obvious advantage over a cloud chamber. Particles traveling through this denser medium are more likely to collide with atomic nuclei and produce interesting developments. A great deal of work has been done to improve nuclear emulsions. In 1947 Pierre Demers of the University of Montreal found a way to prepare stable emulsions containing 90 per cent silver bromide, instead of the usual 30 per cent, and in these more concentrated emulsions particles produce more robust tracks.

Let us proceed to examine some of the identifying tracks. We shall begin by immersing a photographic plate in a very dilute solution of a soluble compound of the radioactive element radium. After leaving it for a time (days, weeks or months) in a dark place, we remove the plate, develop it and inspect it under a microscope. Here and there on the plate we see starlike sets of short heavy tracks, each set radiating like spokes from a hub point. The tracks identify the particles as slow alpha particles, and the formation is known as an alpha star. At the center of the star a radium atom has emitted a series of alpha particles. The radium atom decays first to radon, then to other unstable descendants and finally to lead. In this spontaneous transmutation from radium to lead a total of five alpha particles (plus several beta particles) is emitted. Each in the series comes out with a characteristic kinetic energy, and the different energies (ranging up to 7.7 million electron volts) cause the tracks in a star to be of different lengths.

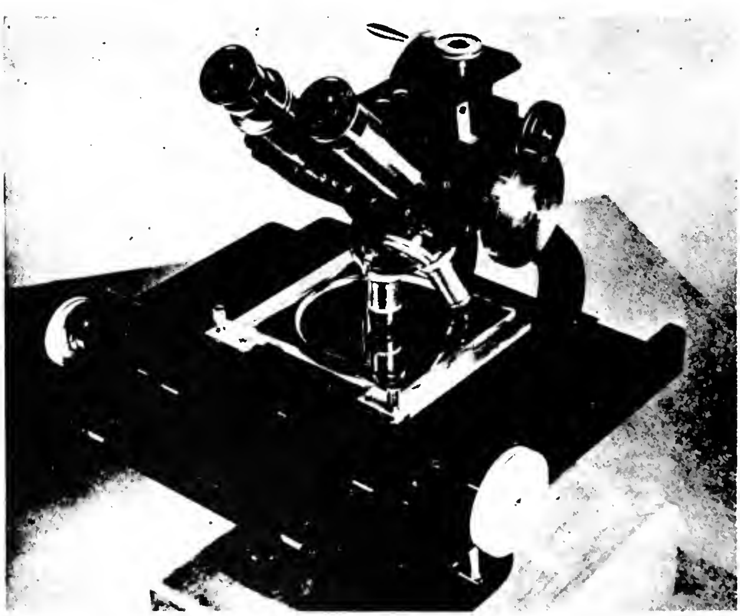
Occasionally the star seen in a photographic plate may represent the disintegration of not one but many radium atoms. This was made clear by an experiment performed by Mlle. C. Chamie at the Curie Institute in Paris. She exposed a plate in an extremely dilute solution of polonium, the last alpha-emitting descendant of radium in the transition to lead. It was supposed that single tracks of alpha particles, from separate atoms of polonium, would appear in the emulsion. Instead Mlle. Chamie found stars consisting of several hundred alpha tracks from a common center. All the

tracks were of the same length, corresponding to the energy of alpha-emission from polonium. Evidently even in an extremely dilute solution the polonium atoms are not completely dissociated into individual ions but may cluster in groups of several thousand atoms. The collections have been named radiocolloids.

All matter contains traces of radioactive substances, and their energy fields have been pulsating in minerals since the earth's crust solidified eons ago. Nature strews the investigator's path with clues—if we could only see. Long before the discovery of radioactivity, geologists had observed that grains in certain minerals, such as mica, were sometimes surrounded with halos of colored material. They could find no way to explain how these colored bands might be formed. In 1907, when radioactivity was a topic of growing interest, John Joly in Ireland noted that the distance from the center of each tiny sphere to the halo around it was about the same as the range of an alpha particle emitted by radium or thorium. He suggested what is now taken to be the correct solution of the mystery: that alpha particles radiating from radioactive atoms at the center ionize iron atoms in the mica near the end of their path, cause the iron to become oxidized and thereby produce the colored bands.

Just as familiar, and as ubiquitous, as the footprints of alpha particles are the footprints of beta particles, or electrons. These light particles make very faint, highly scattered tracks in an emulsion. Originating from radioactive substances and from cosmic ray showers, flying electrons record their presence in emulsions wherever placed or however carefully shielded. Even at great depths underground a photographic plate will show about one million electron tracks per cubic centimeter for each day of its underground exposure.

No footprints are more fascinating than those of the strange particles known as mesons. Had present emulsions been in use in the 1920s, their tracks would have been discovered first and "explained" afterward; as it was, the particles were predicted by the theoretician Hideki Yukawa two years before they were actually found. Yukawa invented the meson to account for the binding force that holds particles together in the atomic nucleus. Tracks of a particle such as he had predicted—about 200 times heavier than the electron—were first discovered in 1937 in cloud chambers monitoring the products of cosmic rays. A mystery soon developed: the theory said that these particles should interact strongly with atomic nuclei, but experiments proved that they were rarely absorbed by nuclei.



SPECIAL MICROSCOPE is used to scan nuclear emulsions. The large stage enables the viewer to follow long tracks. Here the emulsion is a disk embedded in a rectangular Lucite frame fitted with a cover glass. The depth of the track is read on the wheel at upper right.

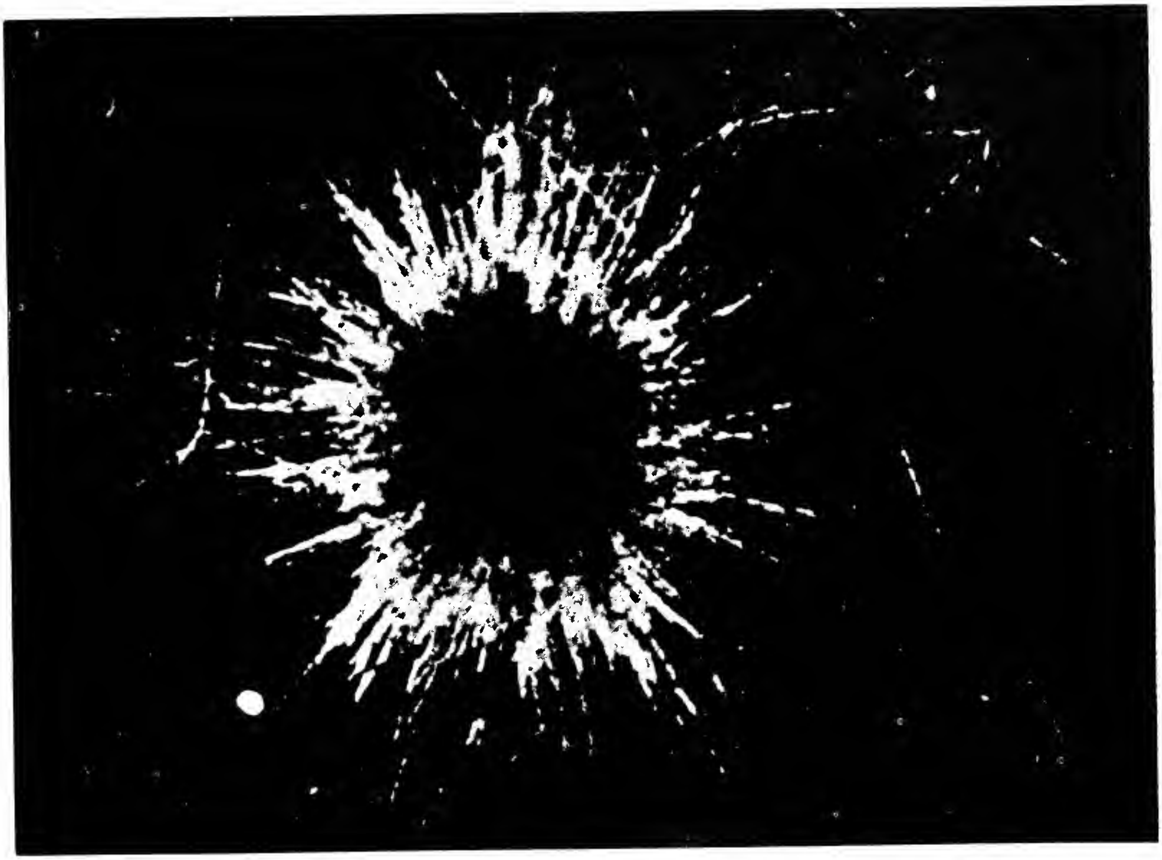
While the theoreticians were pondering this hiatus between theory and experiment, the younger physicists were busy climbing mountains and exposing photographic plates to the intense cosmic radiation high in the atmosphere. By 1947 they had discovered a second, heavier meson which did react strongly with matter [see "The Multiplicity of Particles," by Robert E. Marshak; SCIENTIFIC AMERICAN, January, 1952]. A Bristol University team of investigators headed by C. F. Powell obtained photographs showing that when the heavy pi meson came to rest it promptly decayed into the lighter mu meson.

A year later the young Brazilian C. M. G. Lattes, a member of the Bristol cosmic-ray group, came to the University of California and in cooperation with Eugene Gardner succeeded in detecting mesons from nuclei attacked by a 400-million-electron-volt beam of alpha particles from the Berkeley cyclotron. Two types of pi meson tracks were then identified. Positively charged pi mesons decayed into mu mesons. Negatively charged pi mesons reacted with atomic nuclei, and the disintegration of the capturing nucleus produced a star.

Meanwhile the European investigators, lacking funds for the construction of expensive accelerators, continued to study mesons in the cosmic radiation—the poor man's cyclotron. These simple experiments gave birth to a perplexing number of new particles.

Their first addition to the growing fraternity of Greek-lettered mesons was the tau particle. The Bristol University investigators found its track in an electron-sensitive plate exposed beneath a 12-inch-thick block of lead at the Jungfraujoch High Altitude Research Station. The particle, heavier than a pi meson, produced an unusual three-pronged star on coming to rest. All three prongs could be identified as the tracks of pi mesons. From the available evidence Powell came to the conclusion that the tau meson was an unstable, singly charged particle about 1,000 times heavier than the electron. Powell's brilliant deductions tempt one to finish off his description with the admiring exclamation: "A new particle—elementary, my dear Watson!"

The heavy tau meson is very rare, but an extensive vigil has now detected a number of these events and established the particle's properties. Recent controlled experiments with the six-billion-electron-volt Bevatron at Berkeley indicate that the tau particle and certain other heavy mesons (known as K mes-



ALPHA PARTICLES made the image in this dark-field photomicrograph. The emulsion itself contains tiny colloid particles of radium, one of which is at the center of the image. The tracks were made by alpha particles emitted by radium and its daughter elements.

ALPHA STARS emerged from thorium atoms in this emulsion. The stars at left and right represent the serial decay of single thorium atoms. First the thorium atom emitted an alpha particle, then the daughter isotope emitted another alpha particle, and so on.

ons) probably are all the same particle showing alternate modes of decay.

Neutral particles unfortunately leave no footprints in an emulsion or cloud chamber. They may, however, signal their presence indirectly. For example, a fast neutron charging through an emulsion may collide head on with a hydrogen atom, rip away the latter's electron and cause the proton to recoil and make a track that tells the story of the collision.

At Berkeley all eyes are focused just now on the footprints of the antiproton, which at long last was generated by the Bevatron a few months ago. The antiproton—the negatively charged counterpart of the positive proton—has only a fleeting life, but it makes its existence unmistakably known by the spectacular manner of its death. When the particle comes to rest in an emulsion, there is an explosion which generates a large star.

GROUP	MEMBERS	SYMBOL	REST MASS (ELECTRON MASSES)	MEAN LIFE (SECONDS)
NUCLEONS	PROTON	p^+	1836.13	STABLE
	ANTIPROTON	p^-	1840 ± 90	$\sim 5 \times 10^{-8}$
	NEUTRON	n°	1838.65	750
LEPTONS	ELECTRON	e^-	1	STABLE
	POSITRON	e^+	1	ANNIHILATES
	NEUTRINO	ν	0	
LIGHT MESONS	NEGATIVE PI MESON	π^-	272.8 ± 0.3	2.44×10^{-8}
	POSITIVE PI MESON	π^+	273.3 ± 0.2	2.53×10^{-8}
	NEUTRAL PI MESON	π°	263.7 ± 0.7	5×10^{-15}
	NEGATIVE MU MESON	μ^-	207 ± 0.5	
	POSITIVE MU MESON	μ^+	206.9 ± 0.4	2.15×10^{-6}
HEAVY MESONS	TAU MESON	τ^+	965.5 ± 0.7	$\sim 5 \times 10^{-8}$
	THETA MESON	θ°	965 ± 10	1.6×10^{-10}
	CHI MESON	$\chi\ (K\pi 2)$	963 ± 9	1×10^{-8}
		$(K\mu 2)$	960 ± 7	1×10^{-8}
	KAPPA MESON	$K\ (K\mu 3)$	955 ± 9	1×10^{-8}
		$(Ke3)$	~ 960	
HYPERONS	LAMBDA PARTICLE	Λ°	2182 ± 2	3.7×10^{-10}
	POSITIVE SIGMA PARTICLE	Σ^+	2327 ± 4	$\sim 10^{-10}$
	NEGATIVE SIGMA PARTICLE	Σ^-	2325	$\sim 10^{-10}$
	CASCADE PARTICLE	Ξ^-	2582 ± 10	$\sim 10^{-10}$

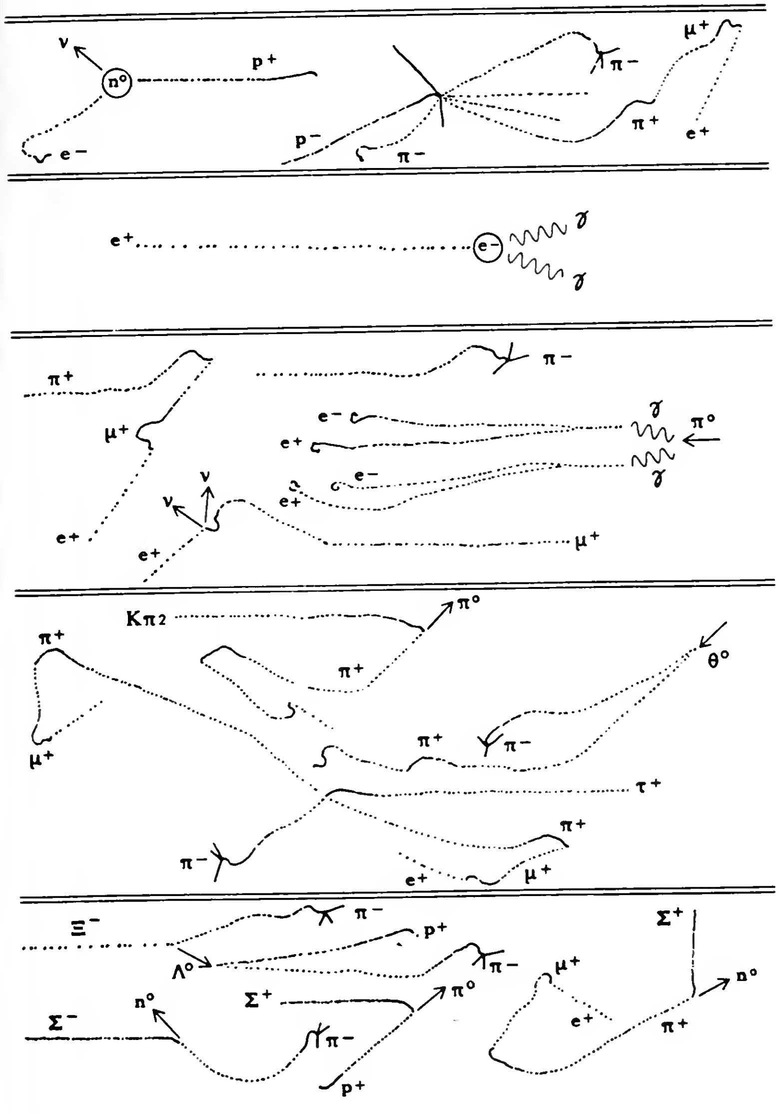
FUNDAMENTAL PARTICLES are listed, together with their characteristic tracks in nuclear emulsions. The photon and graviton are omitted to simplify the organization of the chart. The light mesons are called L particles; the heavy mesons, K particles; the hyperons, Y particles. The chi and kappa mesons have dual symbols, the second of which segregates them according to their mode

The particles emerging from the explosion, among which are several pi mesons, have a large kinetic energy; the total energy released is about that predicted by the theory that the antiproton and a proton combine and annihilate each other, converting mass into energy.

The Bevatron produces antiprotons when a beam of high-energy protons (at 6.2 billion electron volts) hits a copper target. The fast protons attacking the nuclei of the copper atoms generate large numbers of heavy mesons and an occasional antiproton: the yield is about one antiproton per 62,000 mesons. The theory suggests that a high-energy proton interacts with a neutron to form an antiproton-proton pair.

The antiproton has the same mass as a proton. One would therefore expect that it should have about the same probability of collision with atomic nuclei as it travels through matter But experiments with the new particle show that the antiproton actually has about twice as great a collision probability, or cross section, as the proton. This surprising property has presented nuclear physicists with an intriguing problem.

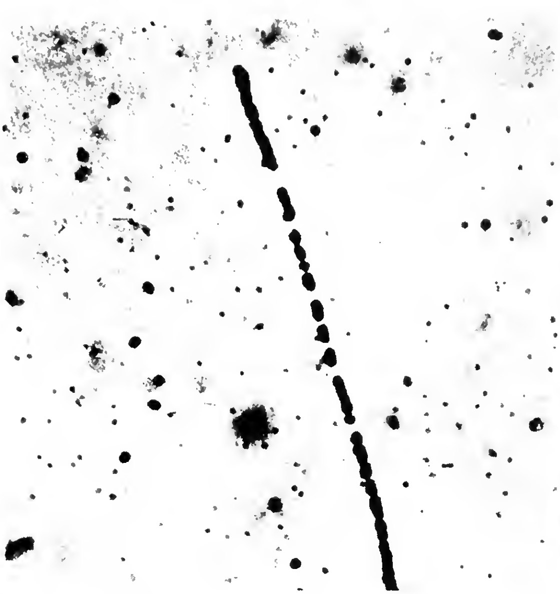
DECAY SCHEME



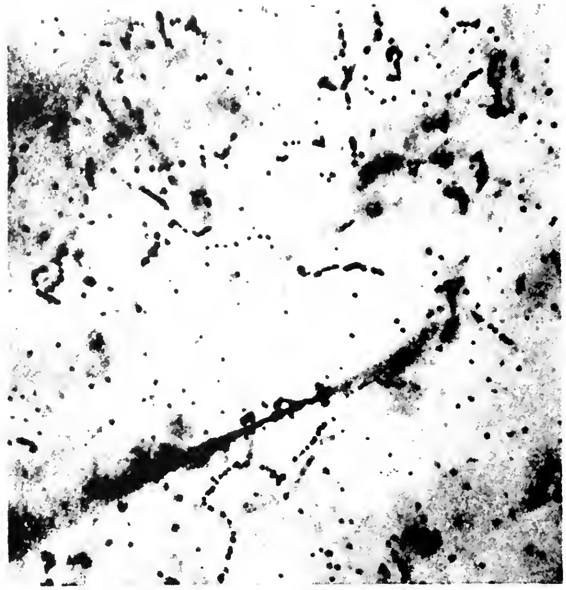
of decay. Kπ2, for example, indicates that this K (not kappa) particle decays into two pi mesons. The decay schemes may be followed by beginning with the particle in that group. The wavy lines (gamma rays), circles and arrows denote particles that do not make tracks.

Enlightening as the work with atom-smashing machines has been, the investigators of particles have not by any means lost interest in the wild assortment of nuclei and nuclear debris that rains into our atmosphere from the bombardment of the cosmic radiation. Of the primary cosmic radiation itself, little reaches ground level, for the atmosphere absorbs it as effectively as would a three-foot-thick layer of lead completely surrounding the earth. But physicists are capturing the footprints of primary particles coming in from space by floating their instruments and photographic plates to the top of the air ocean in balloons. Great impetus was given to this work by the U. S. Navy's development of the plastic "Skyhook" balloon. Unlike rubber balloons, the plastic vehicles can be held at a fixed, preset elevation. Stacks of emulsions have been flown to 100,000 feet—almost at the borders of empty space, for the weight of the overlying air there is only 13 grams per square centimeter, as against 1,030 grams at sea level.

As the primary cosmic rays smash nitrogen and oxygen atoms in the air, they generate a fallout of secondary and tertiary particles. The footprints of these fragments are being recorded at mountaintop stations all over the world. Men who risk their lives to climb a mountain simply "because it is there" are usually very cooperative with the cosmic ray physicists. A light package of photographic plates does not add appreciably to the burden of the climb, and it may add incentive as a form of applied mountaineering. In the ascent of Mt. Everest Sir Edmund Hillary took a small package of plates (given him by Professor Eugster of Zurich University) to the 25,850-foot camp site. Unfortunately, in the excitement of the triumphant descent from the peak the plates were overlooked. Sir John Hunt, the leader of the expedition, apologized in his book, *The Conquest of Everest:* "I very much



SLOW NEUTRON gave rise to this track in an emulsion containing lithium borate. The neutron encountered a lithium atom at the lower end of the short, heavy line at the top. The track was then made by two fragments of the nucleus recoiling from each other.



ELECTRONS made the faint, wavy tracks in this emulsion, which was aged for 50 days before it was developed. The heavy track at the bottom was made by an oxygen nucleus in primary cosmic radiation. The electron tracks along this image are called delta rays.

regret to say that the plates have remained on the South Col, where they must by now have made a very definite recording of . . . cosmic ray phenomena."

Among the first to get a recording of these phenomena was Marietta Blau of the University of Vienna. Nineteen years ago she exposed a series of photographic plates for four months on a mountaintop at Innsbruck. When she developed them, she found not only the familiar alpha stars from radioactive substances but also a number of bigger stars with much longer, less dense prongs. The tracks evidently were produced chiefly by protons. Dr. Blau surmised correctly that they were the debris of nuclei disrupted by cosmic rays; she followed up this finding and today is studying nuclear disruptions produced by the Cosmotron at the Brookhaven National Laboratory.

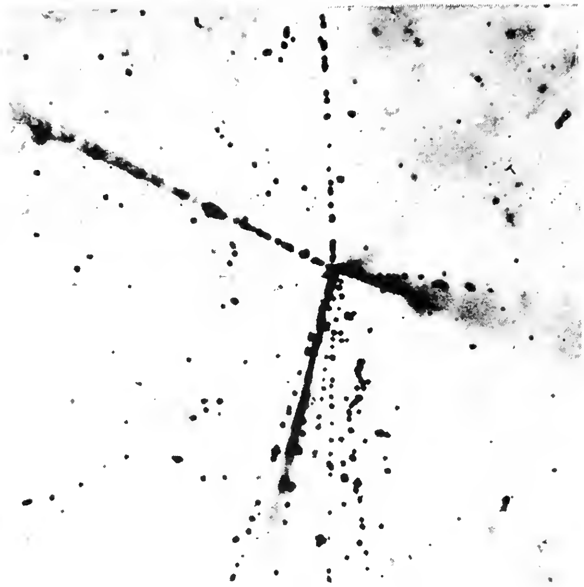
The smashing of nuclei by cosmic rays increases rapidly with altitude. At sea level in northern latitudes the rate of star production in photographic plates is about one per cubic centimeter of emulsion per day of exposure; at 14,260 feet on Mt. Evans in Colorado the rate is 20 times that; and in balloons near the top of the atmosphere, 2,500 times.

The tracks of the primary cosmic particles that arrive there from space are often extremely robust. These thick tracks are made by heavy nuclei, much larger than the nuclei of hydrogen atoms. The track is covered with a fur of spurs projecting from its sides—secondary ionizations which are known as delta rays. Since the amount of ionization by a particle along its path is proportional to the square of its charge, the amount of delta-ray ionization identifies the particle. The primary cosmic particles have been found to include the nuclei of almost all the elements from hydrogen to nickel. Iron nuclei often produce tracks heavy enough to be seen with the naked eye.

Sometimes the incoming heavy nu-

IRON NUCLEUS in primary cosmic radiation entered this picture from the left. Escaping catastrophic collision with nuclei in the emulsion, it finally came to rest at the right. Its energy was dissipated by a series of encounters in which it removed electrons

NEGATIVE PI MESON made the track between these two stars. At the top is a nucleus disrupted by a primary cosmic ray. At the bottom is a second nucleus disrupted by the pi meson. Negative mesons are readily absorbed by nuclei because of their opposite charge.



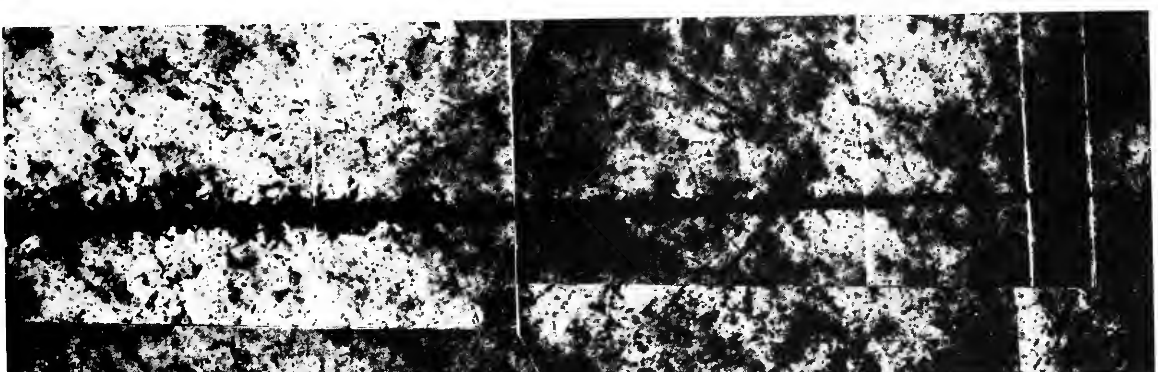
PROTON in primary cosmic radiation made the nearly vertical track at the top of this emulsion. The tracks produced by its encounter with a nucleus in the center of the emulsion are characteristic of fragments and/or particles with a single electric charge.

cleus is partly sheared off by a glancing collision in the air, and the separated bundles of nucleons diverge from the point of collision. Sometimes the cosmic primary hits an atom head on and disintegrates it, emitting a shower of heavy mesons: as many as 200 charged mesons have been seen in a single star. Many of the pi mesons decay during flight into mu mesons; the latter, nearly immune to capture by atoms, zip through the atmosphere and often plunge deep into the earth.

A small proportion of the heavy nuclei from space escape catastrophic collisions and are eventually slowed down by ionization processes in the atmosphere. When these particles are caught in an emulsion, they produce very spectacular tracks. The track is first thick and furry; then as the heavy nucleus slows down and begins to pick up electrons, the reduction of its positive charge diminishes the ionization it produces, so that its track tapers down to a needle point at the end of its flight.

The last grain at the rest point of a heavy primary cosmic particle is a thing to marvel at. Embedded within the grain of silver in the emulsion is an atom with a history unlike that of its neighbors. It is an atom which may have been blown out of a star in our galaxy millions of years ago. It was accelerated through interstellar space by magnetohydrodynamic fields. For millions of years it escaped collision with cosmic dust. Finally it plowed into the earth's atmosphere, and in a single moment lost its store of energy accumulated since birth. Such is the ever-increasing entropy of the universe, of which Swinburne wrote:

We thank with brief thanksgiving
Whatever gods may be
That no man lives forever,
That dead men rise up never;
That even the weariest river
Winds somewhere safe to sea.



from atoms in the emulsion. These electrons made the wavy tracks along the path of the iron nucleus. The track is about a 16th of an inch in length, too long to be shown in a single photomicrograph. It has accordingly been depicted in a mosaic of photomicrographs

Our knowledge of elementary particles depends on the spark chamber and similar devices which make visible the tracks of these subatomic particles.

6 The Spark Chamber

Gerard K. O'Neill

Scientific American article, published in 1962.

The present understanding, imperfect but growing, of the fundamental nature of matter has come largely from observation of the elementary particles. The protons, neutrons, electrons, mesons and other particles reveal the most when they can be studied one at a time or when only two or three of them interact. When larger numbers are present, the sheer mathematical complexity of their interaction hides the fundamental simplicities. For this reason the efforts of many experimental physicists over several decades have gone into the development of sensitive methods for detecting single particles.

There is no single best design for a particle detector. To obtain certain characteristics it is usually necessary to sacrifice others, and the choice depends on the nature of the experimental "events" one wishes to observe. Physicists working with the large particle-accelerating machines have increasingly been concerned with extremely rare events, epitomized by the recent discovery at the Brookhaven National Laboratory that there are two kinds of neutrino rather than one [see "Science and the Citizen," page 52]. To obtain the evidence for this discovery the 30-billion-electron-volt proton accelerator at Brookhaven was operated for six months. Over this period the number of recorded events caused by neutrinos averaged fewer than one every three days. The particle detector used in the experiment is of an entirely new type: it is called a spark chamber. Before explaining its operation I shall describe the general nature of the particle-detection problem.

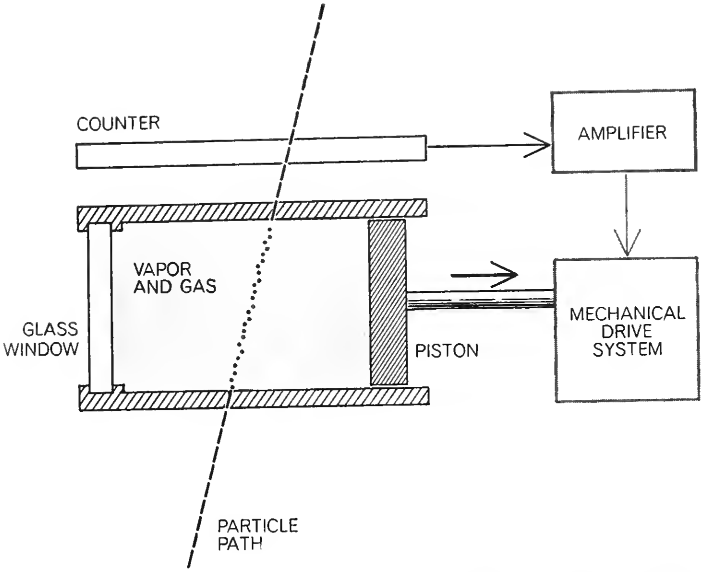
The problem is far from easy, because an elementary particle can pass freely through many atoms of any substance without leaving a trace. Even at present there is no practical device that can detect electrically neutral particles without destroying or deflecting them. Charged particles, however, exert a strong electrostatic force on the electrons of the atoms through which they pass. Usually the electrostatic force between the negative electron and the positive nucleus is enough to keep the electrons from breaking free, but occasionally—roughly once in every 1,000 atoms through which a charged particle passes—an electron is jolted loose. In air, for example, about 100 electrons are freed along each centimeter of the path of a charged particle, and for each free electron a corresponding positive ion is formed. If the small amount of energy contained in this "ionization trail" can be made to produce some visible effect, the physicist can find out where the particle went. He can also measure the momentum of a particle by observing the radius of curvature of its track in a magnetic field, and he can obtain information about the way it interacts with other particles by observing sudden changes in direction of its track.

In one of the first of all elementary-particle experiments Hans Geiger and Ernest Marsden, working in the Cavendish Laboratory at the University of Cambridge, detected the small energy of an ionization trail without amplification by using the extreme sensitivity of the dark-adapted human eye. They observed the small flashes of light made when alpha particles went through certain crystalline materials called scintillators. From Geiger and Marsden's observations of the angles at which alpha particles scattered from a target into the scintillator, Ernest Rutherford concluded by 1913 that the positive charge of the atom was concentrated in a nucleus.

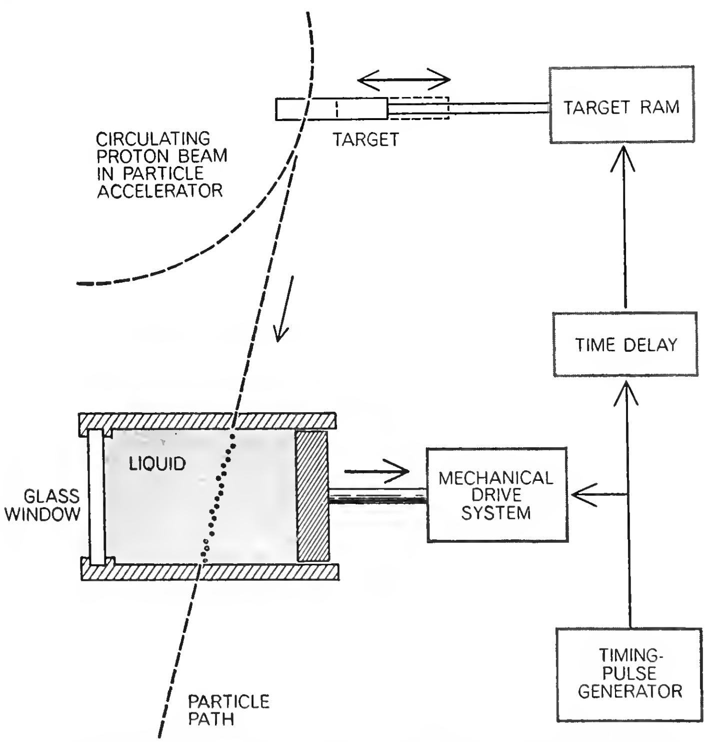
A fast, singly charged particle—a cosmic ray meson, for example—produces only about a thousandth as many free electrons per millimeter of track as a slow, doubly charged alpha particle does. The detection of fast particles therefore requires some kind of amplification of the energy of the ionization trail. Since Rutherford's time the devices used to detect elementary particles have divided into two broad classes, both of which amplify. One class consists of "counters." Every counter includes a sensitive volume of gas, liquid or solid with well-defined dimensions in space. When a charged particle passes through the sensitive volume, the counter produces a brief electric pulse, or signal. The pulses can be tallied electronically; hence the name "counter."

The other class does not have a well-recognized generic name, but it can be called the class of "track detectors." A track detector shows where a charged particle went by indicating many points in space along the particle's ionization trail. Usually the information provided by a track detector is recorded by photography. In fact, for certain purposes stacks of photographic film or a single block of photographic emulsion can be used directly as a track detector. A charged particle sensitizes emulsion grains along its track and amplification is achieved by means of a chemical developer. In the next few years some advanced track detectors may be built that will put out information in the form of electrical signals.

If one compares the two classes, it is apparent that the counter gives only a limited amount of information, but it gives it immediately in a simple form suitable for direct use in electronic circuits. In modern counters the information is often available in less than 10 nanoseconds (10 billionths of a second). The track detector gives much more information, but the information goes into photographic emulsion, where it is unavailable until the emulsion is developed



CLOUD CHAMBER, invented in 1911 by C. T. R. Wilson, was the first of the particle-track detectors. A counter, which simply senses the arrival of a particle, triggers the movement of a piston that expands the gas and vapor inside the chamber. This makes the vapor supersaturated, and fog droplets rapidly grow along the ionization trail left by passage of the particle. The droplets form clear tracks, which are photographed stereoscopically for analysis.



BUBBLE CHAMBER, a track detector invented by Donald A. Glaser, contains a liquid near its boiling point. When the chamber pressure is lowered, the liquid becomes superheated and bubbles of vapor grow along the ionization trail left by a charged particle. A timing mechanism moves a target into the beam of circulating protons in an accelerator, thereby directing particles into the chamber at the instant it is most sensitive to bubble growth.

and analyzed. A counter with a sensitive volume of a cubic foot can only signal that a charged particle has passed somewhere within that cubic foot. Some track detectors with the same sensitive volume can indicate each point of the particle's path within a thousandth of a centimeter. The space resolution of the track detector balances against the reporting speed of the counter.

In modern elementary-particle experiments the experimenter often wants to trace all or part of the life histories of particles entering his detectors. He wants to identify the mass, charge and frequently the energy of each particle that enters. In addition he wants to observe if and in what way the entering particles react with the atoms in his detector. If new particles are produced by reactions, he wants to measure the properties of these product particles and to see if they decay spontaneously into combinations of other particles. In most cases, the rarer the reaction, the greater its significance. Typically only one in many thousands of particles entering a detector will produce an interesting event. If the experimenter's apparatus includes track detectors, it is much to his advantage to use counters to select those events that are worth recording in the track detector. Otherwise he may have to search through hundreds of thousands of pictures to find the rare events of interest.

The first successful track detector was the cloud chamber, invented by C. T. R. Wilson in 1911. Wilson recognized that a supersaturated vapor is unstable and that the vapor will condense into droplets around any available free ions. In cloud chambers (which are still used) a saturated vapor is maintained in a closed volume under well-controlled conditions of temperature and pressure. When a charged particle passes through the chamber, the ionization trail it leaves persists for a fraction of a second. Either before or directly after passing through the cloud chamber the particle traverses counters, which produce an electric pulse. The pulse, signaling the passage of a particle, is made to initiate the outward motion of a piston; this allows the gas inside the chamber to expand and renders the vapor in the gas supersaturated [*see top illustration at left*]. The vapor then begins to form droplets of fog, which condense around the ions of the charged-particle track. Droplets also tend to form around dust particles or droplets left over from a previous expansion. But under the right conditions (achieving

them is rather tricky) there forms in the chamber, in a fraction of a second, a clear trail of vapor droplets, which shows with good fidelity the path of the particle that triggered the counters. The advantage of the cloud chamber is that it can be triggered. A chamber may remain idle for hours waiting for a rare cosmic ray event, but when the event occurs and is recognized by the counters, the chamber operates on demand to record it.

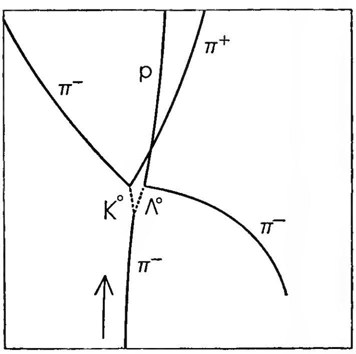
Unfortunately cloud chambers have two rather serious drawbacks. First, the device is slow to set in operation, and the ionization trails persist for a large fraction of a second. As a result the number of incoming particles must be limited to prevent chamber pictures from being cluttered with more tracks than one can "read." The second drawback is the difficulty of putting into the chamber materials with which one might like to see particles interact. If material is introduced in the form of plates, the plates must be relatively few and widely spaced; otherwise the chamber will not work. If much material is needed, it must therefore be in the form of thick plates, with the result that interactions tend to occur deep in the plates, where the tracks cannot be seen. It is rather like Greek tragedy, in which all the mayhem occurs offstage and the audience is treated only to a secondhand account of it.

In the early 1950's Donald A. Glaser, then at the University of Michigan, developed a new type of track detector, the bubble chamber, for which he received a Nobel prize in 1960. This detector is also based on an amplification principle—the growth of bubbles in a superheated liquid. Some of the energy from an ionization trail goes into a few fast electrons, which can give up 1,000 or 2,000 volts of energy in a small volume to produce rapid local heating. If the trail is in a liquid that has suddenly been superheated by expansion, the bubbles will tend to grow fastest along the "heat track" and only slowly in other parts of the liquid. Glaser's invention was soon in use in many laboratories throughout the world, and it is safe to say that by 1959 more than half of all experimental research in elementary particle physics employed the bubble chamber.

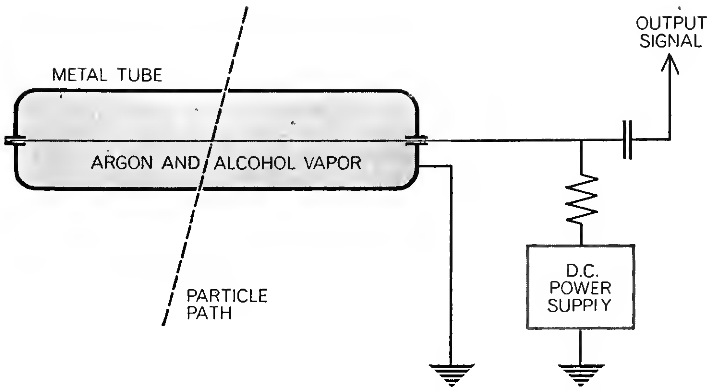
An important virtue of Glaser's device is that one can fill the chamber with a wide variety of liquids, choosing the one that provides interactions of particular interest. For many purposes liquid hydrogen is ideal because it presents as a target for incoming particles only electrons and protons. In all other substances neutrons are also present. Other useful liquids are propane—in which the target atoms are carbon and hydrogen—and xenon, whose massive nucleus (54 protons and 77 neutrons) provides high stopping power. In addition the bubble chamber produces particle tracks of higher definition than those made by any other track detector, except for tracks made directly in photographic emulsion.

The bubble chamber shares with the emulsion method one serious disadvantage: it cannot be triggered. Since there is no way to select rare events one has no choice but to photograph the chamber at every expansion cycle, develop the films and examine hundreds or thousands of exposures looking for events of interest. Triggering is impossible because the heat track produced by a charged particle cools down in much less than a millionth of a second. This is far too short a time for the mechanical expansion system to set the chamber in operation. As a result bubble chambers are used almost exclusively with large accelerators, where a timing se-

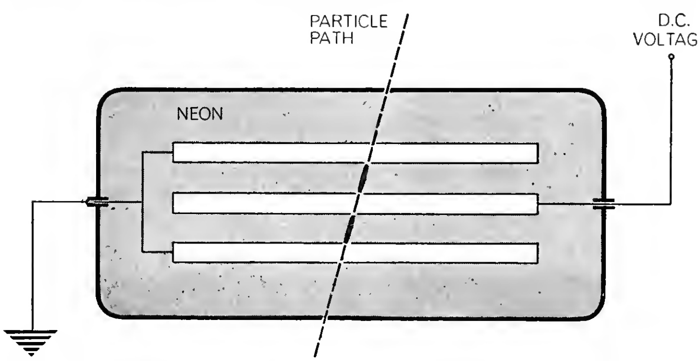
p PROTON
π^+ POSITIVE PI MESON
π^- NEGATIVE PI MESON
Λ° NEUTRAL LAMBDA PARTICLE
K° NEUTRAL K MESON
e^+ POSITRON
e^- ELECTRON
γ GAMMA RAY



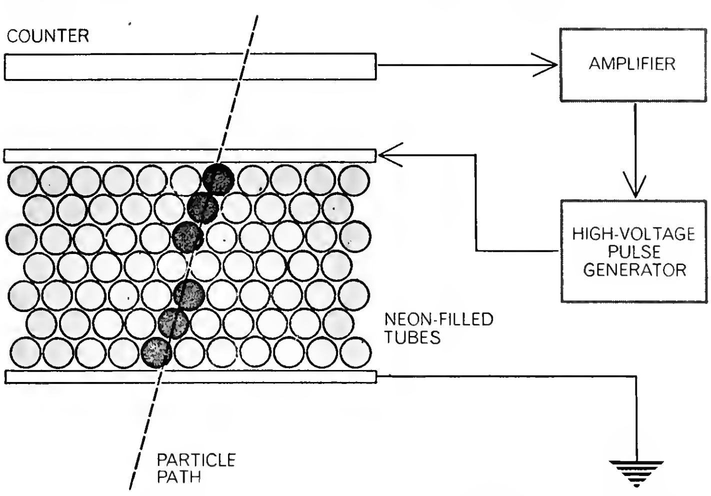
BUBBLE CHAMBER TRACKS ***(right)*** **were photographed in the 72-inch liquid-hydrogen bubble chamber at the Lawrence Radiation Laboratory of the University of California. The map and key at left identify the particles taking part in the event recorded.**



GEIGER-MÜLLER COUNTER, invented in 1928, was the first device to use the amplification process available in an electric spark to detect the passage of a charged particle. A central wire inside a tube is placed at high voltage. Electrons set free from gas atoms by the passage of a particle are accelerated by the strong electric field and free other electrons in a chain reaction. The result is a large output pulse that needs no amplification to be detectable.



SPARK COUNTER was a nontriggered forerunner of the spark chamber. A high constant voltage is maintained on a metal plate placed between two grounded plates. Passage of a charged particle provides free electrons that initiate sparks in the gas between the plates.



HODOSCOPE CHAMBER, another forerunner of the spark chamber, utilizes the triggering scheme usually employed with cloud chambers. The chamber consists of neon-filled glass tubes stacked between two metal plates. When a charged particle trips the counter, a high-voltage pulse is sent to the plates, placing the tubes in a strong electric field. Tubes through which the particle passed contain ions and free electrons and therefore glow.

quence first expands the chamber, then sends in a burst of particles to be analyzed [*see bottom illustration on page 38*]. The chamber must then be given about a second in which to recover.

Unlike the cloud chamber and the bubble chamber, the spark chamber was the work of many hands. Its development was based on one of the most spectacular methods known for making ionization trails visible—the electric spark. The generation of an electric spark is an extremely complicated process, but it is clear that under some conditions a spark can develop from a type of chain reaction. The reaction starts when an electron from an ionized atom, accelerated by a strong electric field, bumps into and ionizes other atoms. The electrons from these atoms cause further ionizations, leading in a very brief time to a brilliant electric spark. In 1928 the amplification process available in the electric spark was used in the first of all electrical detectors for single charged particles, the Geiger-Müller counter. In this simple device, named for Hans Geiger and Walther Müller, a central wire inside a tube is charged to high voltage. When a particle goes through the counter, the electrons of its ionization track are swept toward the wire. Accelerating as they approach the wire's strong field, they ionize more atoms. The ionized atoms emit photons (light quanta), which release additional electrons from the gas, spreading the discharge. Within millionths of a second the gas all along the center wire serves as the path for an electric spark. Geiger counters make tremendous pulses, which was a great virtue when sensitive electronic amplifiers were still difficult to build.

In the 1930's the standard equipment of the elementary-particle physicist consisted of a cloud chamber triggered by Geiger counters. In the late 1940's, when Geiger counters had been generally superseded by the development of scintillation counters (faster and capable of giving more information), a few physicists began trying to use the mechanism of the electric spark in a detector that would make visible the track—not just the presence—of a charged particle. J. W. Keuffel, working at the California Institute of Technology and later at Princeton University, built several spark counters, consisting of well-polished condenser plates kept at high voltage. If the plates were carefully aligned, clean and dust-free, and maintained just below the potential needed for a spark to jump between them, they would sometimes spark preferentially along the trail

of an incoming cosmic ray particle. Keuffel suggested the use of arrays of his parallel-plate spark counters to obtain tracks of the passage of a charged particle, but these counters were so difficult to build and to operate that it was not easy to follow up the suggestion.

In 1955 M. Conversi and A. Gozzini described in the Italian physics journal *Nuovo Cimento* an intermediate type of track chamber somewhat similar to the Keuffel spark counter. Their device, called a hodoscope chamber, consisted of many neon-filled glass tubes stacked between two parallel metal plates [*see bottom illustration on opposite page*]. Within a few millionths of a second after the passage of a charged particle through the stack of tubes, a set of counters outside the stack triggered an electronic circuit that placed a strong electric field on the tubes. Those through which the particle had passed then glowed, much as a neon sign glows. Other tubes remained dark if the applied pulse was on for only a short time. The hodoscope chamber was fairly easy to build, and its inventors had introduced a technique that was essential for the development of spark chambers: the use of counters to pulse the electric field. In their chamber the high voltage was on only when they were sure a particle track was there to be photographed. If the high voltage had been left on continuously, as it was in the earlier spark counters, some neon tubes would eventually have fired even in the absence of an entering track. The chief defect of the hodoscope was that it revealed only two dimensions of a particle's three-dimensional path.

In 1957 two British physicists, T. E. Cranshaw and J. F. de Beer, reported in *Nuovo Cimento* the next step toward a practical spark chamber. They combined the parallel-plate geometry of the spark counter with the pulse-triggering technique of the hodoscope chamber to make an efficient spark chamber with six one-millimeter gaps. They also introduced the use of a continuous electric clearing field to remove from the chamber ionization trails older than a few microseconds. This electric field, well below the threshold needed to make a spark, caused a slow continuous drift to the plates of all free electrons and ions released in the chamber gas. In this way it "erased" ionization trails in a few microseconds. A similar clearing field had long been used in cloud chambers to sweep out the slow-moving positive ions.

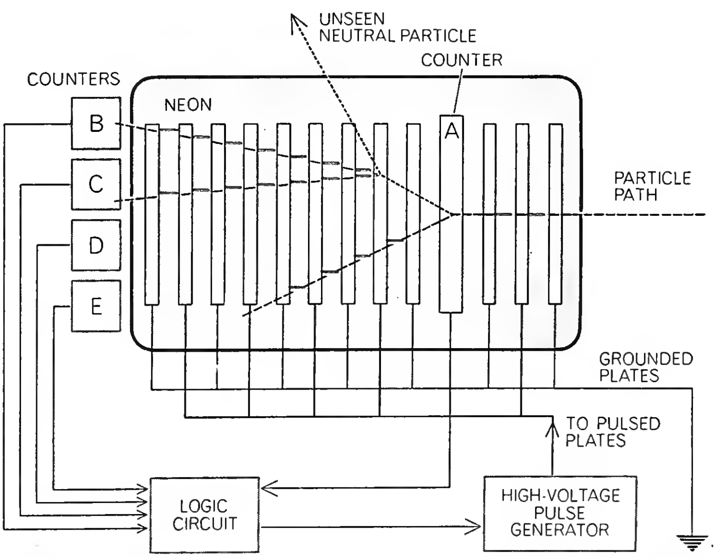
It happened that Cranshaw and de Beer chose to use air rather than neon in their chamber, and this small difference made it impossible for their chamber to detect two or more simultaneous tracks. Still, their work was so successful that several other groups—in Germany, Japan, the U.S.S.R. and the U.S.—continued to work along similar lines.

The final step—substitution of neon for air—was taken by S. Fukui and S. Miyamoto of Osaka University and reported in 1959. The two Japanese physicists were interested in developing a track detector that could be used for cosmic rays. Bubble chambers are not useful for such work, since they cannot be triggered. Fukui and Miyamoto found that in a chamber containing neon rather than air several simultaneous particle tracks could be seen.

One big difference between the behavior of air and of neon in spark chambers is that oxygen molecules (O_2) in air can combine with the free electrons of the ionization trail, whereas neon atoms cannot. The inertness of neon—and of other "noble" gases—is explained by the fact that it has a full complement of eight electrons in its outer electron shell. In contrast an oxygen molecule can acquire one electron and thereby become a negative ion (O_2^-). The electrons are well anchored to the oxygen molecules, some 60,000 times more massive than themselves, and cannot be freed except by application of a strong electric field. Consequently an air-filled spark chamber requires an operating pulse of 7,000 to 10,000 volts for each millimeter of space between its plates. This is about three times the voltage needed for a neon spark chamber.

The formation of oxygen ions also explains other characteristics of an air spark chamber. If the electron in an ionization trail can migrate freely to the plates of the chamber, its travel time is brief. But if it is attached to an oxygen molecule along the way, the velocity of the resulting ion is much slower than that of the electron. In fact, if the mass of a particle is suddenly increased by 60,000 times, its velocity must decrease by the square root of 60,000, or by a factor of about 250. Because most of the electrons liberated in an air spark chamber are slowed down in this fashion, they require many microseconds to migrate to the plates of the chamber. Such a chamber therefore remains sensitive for a long time, and in it old tracks cannot be quickly erased.

It is not so clear why air chambers show only one spark per gap even though several ionization trails may be present. It may be that at the high electric fields needed to operate such chambers the spark produced by the first electron freed from an oxygen ion occurs so rapidly that the plates are quickly dis-

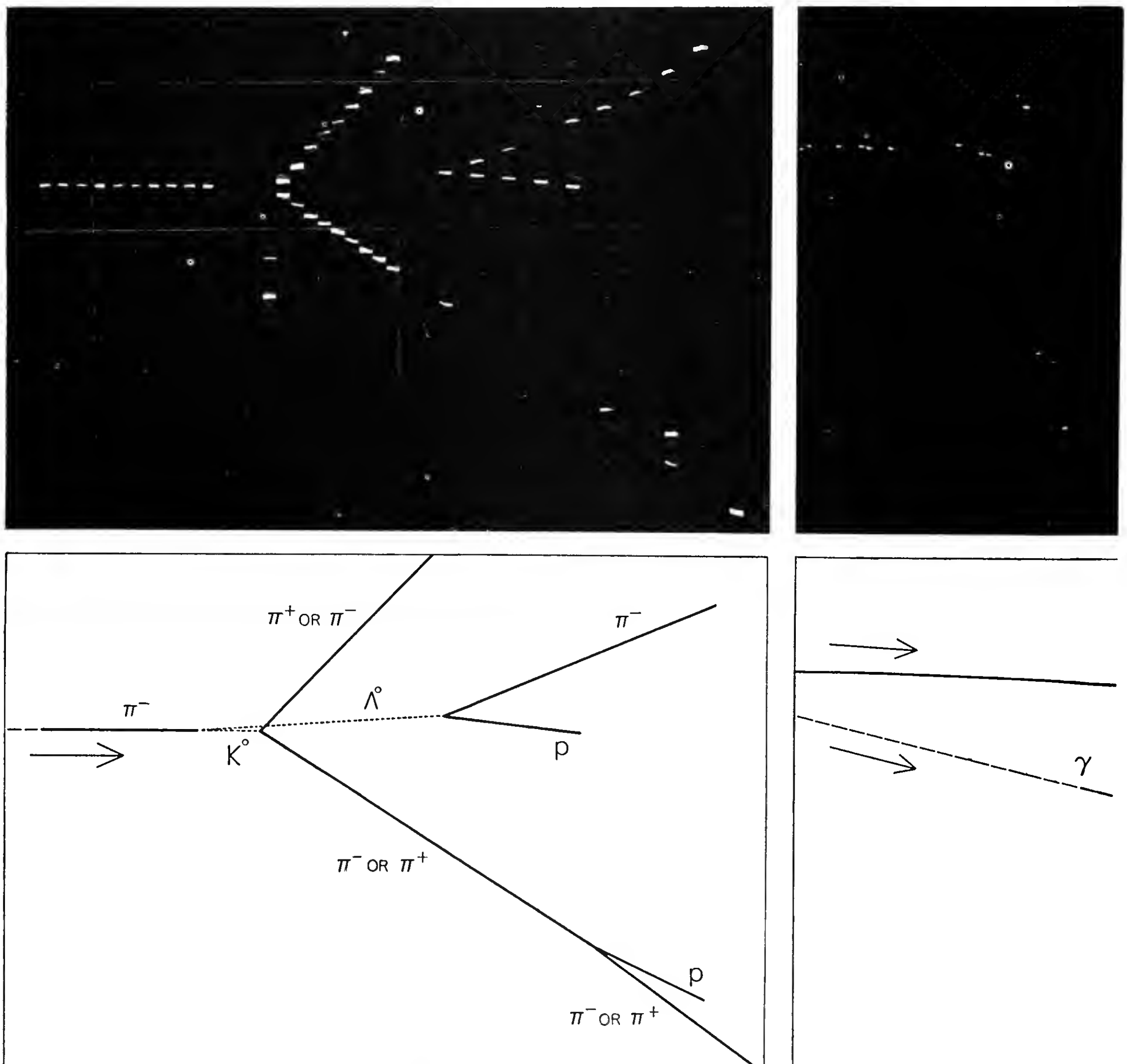


SPARK CHAMBER, which became practical with the work of S. Fukui and S. Miyamoto in 1959, consists of an array of thin metal plates surrounded by neon. It is also provided with counters and a "logic" circuit for determining when a particle meeting certain criteria has appeared. When it appears, a high-voltage pulse is sent to alternate plates and sparks occur along the ionization trails left by each charged particle. In the example shown, a charged particle interacts in counter *A*, yielding one neutral and one charged secondary. The secondary decays in the chamber, producing two charged particles and a neutral one.

charged below the threshold field, preventing any other attached electrons from getting free to start other sparks. This is consistent with an observation by Cranshaw and de Beer that only one electron is needed to start the spark.

Following the announcement of a practical spark chamber by Fukui and Miyamoto in 1959, the idea was immediately taken up by physicists in the U.S. and elsewhere. Within a matter of months Bruce Cork of the University of California had built a six-gap spark chamber and had operated it in a beam of particles from the six-billion-electron-volt accelerator of the Lawrence Radiation Laboratory. Almost simultaneously James L. Cronin of Princeton University built and operated a large 18-gap spark chamber, which yielded high-quality pictures of the tracks made by cosmic rays and by accelerator-produced particles. Both of these chambers used noble gases (neon or argon) and employed clearing fields to erase the ionization trails. Cork and Cronin were also the first to conduct actual experiments using a spark chamber as a particle detector.

In their work, as in most subsequent experiments using spark chambers, the occurrence of an interesting event was recognized by a system of conventional counters, which then triggered the operation of the chamber. Typically particles arrived at the spark chamber at intervals of a few microseconds and their tracks were swept to the plates by the continuous clearing field after only one microsecond. Consequently the pulsing of the spark chamber had to be carried out in much less than one microsecond so that

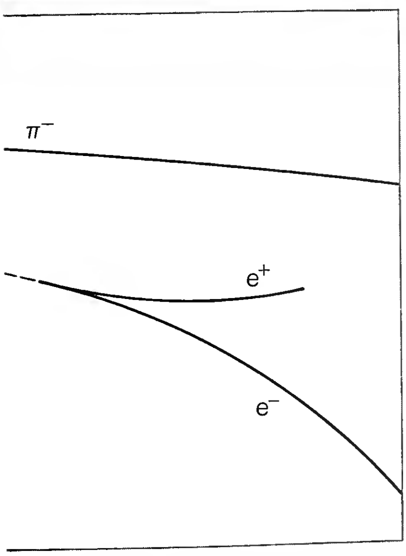


SPARK CHAMBER PICTURES show the appearance of particle tracks when the particles are curved by a magnetic field (*top right*) and when they are not (*top left*). The maps below each picture identify the charged particles, which leave tracks, and the neutral ones, whose presence is inferred. The reaction at the left was seen in a spark chamber operated at Brookhaven by James L. Cronin

the interesting track would still be there to be detected by spark amplification.

Within the past three years a wide variety of spark chambers have been built, each designed to exploit certain desirable features. Some have been made with thick carbon plates to allow interactions of the incoming particles with carbon. Others have been built in the form of a cylinder, to study the scattering of particles by a target located on the axis of the cylinder.

Along with several other physicists, I have been particularly interested in the design and use of thin-plate spark chambers that can be operated in a magnetic field. In a uniform magnetic field the path of a charged particle of constant energy is a circle whose radius is proportional to the momentum of the particle. The idea of using a magnetic field to obtain momentum information goes back to the early days of the cloud chamber, and bubble chambers are nearly always operated in such a field. The measurement of the momentum of each charged particle in a reaction is always useful, and frequently essential, for identifying the particles and learning the details of their interactions.



of Princeton. The picture at right was made in author's two-cubic-foot spark cnamber at Brookhaven, shown at bottom of page 36.

When a magnetic field is used in a spark chamber, the sparks trace the ionization trails more closely if the spacing between the chamber plates is small. As the spacing is reduced, however, it becomes increasingly important for the plates to be flat and uniformly spaced, and the triggering pulse has to rise from zero to the peak voltage at higher speed. Fukui and Miyamoto had used spacings of 10 millimeters. Cork's chamber had a six-millimeter spacing. Within a few months we found in our laboratory at Princeton University that the spacing between spark-chamber plates operated in neon could be as small as two millimeters.

Unless very close plate-spacing is wanted, the construction of a spark chamber is not too difficult and might make a feasible project for an amateur scientist. A chamber with an adjustable plate spacing of two to 10 millimeters, the first model built by our group, was largely the work of college sophomores majoring in physics. Our second instrument was small but operated in a magnetic field. It contained 50 gaps of three millimeters each, separated by aluminum foil a thousandth of an inch thick. A third chamber, with 128 gaps of three-millimeter spacing and a volume of two cubic feet, can measure the momentum of particles with good accuracy. When the tracks cross 100 or more gaps, the accuracy of momentum measurement approaches that obtainable in a good bubble chamber.

At present the advantages the bubble chamber retains over the spark chamber are two. First, pure liquid hydrogen can be used as the only material in the bubble chamber, thereby limiting nuclear reactions to those between elementary particles and hydrogen nuclei (protons). In 1960 we studied the possibility of imitating a hydrogen bubble chamber by using liquid-hydrogen-filled hollow plates in an atmosphere of gaseous helium. We established that such a chamber would work but so far no one has needed its properties badly enough to build one. The second advantage of the bubble chamber is that it yields very fine ionization trails, and it produces them no matter which way the particle is moving. The bubbles trace a particle's path with an uncertainty of less than a thousandth of an inch. Even in narrow-gap spark chambers the sparks scatter in a region 15 or 20 thousandths of an inch wide. Moreover, in a spark chamber the path uncertainty increases as the particle approaches a course parallel to the plates.

In spite of these drawbacks the spark chamber has two big advantages over the bubble chamber. First, the decision to photograph a given event can be made after the event has occurred. Second, because old ionization trails are swept to the walls after only one or two microseconds the spark chamber picture shows only the tracks produced during the last microsecond before the chamber was pulsed. Because of these two features one can select and photograph an interesting event caused by a single entering particle out of many thousands, all arriving over a few thousandths of a second. Each ionization trail of the uninteresting majority of tracks is swept away and does not remain to confuse the picture.

The decision as to which events to photograph is made by "logic" circuits that analyze the output of counters, which may be located outside or inside the spark chamber itself. Frequently the logic requirements are severe and the pulses from many counters must be digested and analyzed before a decision is made whether to pulse the chamber or not. Ordinarily a time of about 100 nanoseconds (100 billionths of a second) is available for the decision. This is not uncomfortably short with present-day circuitry. For the past 10 years it has been practical to use circuits that operate in 20 nanoseconds or less.

Those of us who have jumped on the spark chamber bandwagon are naturally enthusiastic about future prospects for the instrument. We have found that physicists who formerly used bubble chambers are delighted to have a device that eliminates great masses of uninteresting pictures. And former counter physicists are happy to see the tracks they knew were going through their counters. We all know that neither bubble chambers nor counters are going to be put out of business by the new track detectors, but to a remarkable degree spark chambers allow us some of the best of both worlds.

This speech is a lucid historical introduction to the cyclotron, with frank references to missed opportunities.

7 The Evolution of the Cyclotron

Ernest O. Lawrence

Nobel Prize lecture given in December 1951.

The development of the cyclotron was begun more than twenty years ago and perhaps it is appropriate on this occasion to give something of an historical account. The story goes back to 1928 when I had the good fortune of becoming a member of the faculty of the University of California. At that time it seemed opportune to review my plans for research, to see whether I might not profitably go into nuclear research, for the pioneer work of RUTHERFORD and his school had clearly indicated that the next great frontier for the experimental physicist was surely the atomic nucleus.

It seemed equally obvious also at that time that a prerequisite to a successful experimental attack on the nucleus was the development of means of accelerating charged particles to high velocities — to energies measured in millions of electron volts, a task which appeared formidable indeed! Accordingly, I devoted considerable time and thought to the technical problem of ways and means of reaching millions of electron volts in the laboratory. The problem seemed to reduce itself to two parts, *A* the production of high voltages and *B* the development of accelerating tubes capable of withstanding such high voltages.

Since transformers and rectifiers for such high voltages seemed rather out of the question for various reasons, not the least of which were connected with financial limitations, I naturally looked for alternative means of producing high voltages — the surge generator which was used by BRASCH and LANGE — the electrostatic generator which Professor W. F. G. SWANN was working on when I was a student under him at the University of Minnesota in 1924 and which was later brought to practical development by VAN DE GRAAFF, and the Tesla coil source of high voltage which TUVE, BREIT and HAFSTAD brought to a fruitful stage of development.

One evening early in 1929 as I was glancing over current periodicals in the University library, I came across an article in a German electrical engineering journal by WIDEROE on the multiple acceleration of positive ions. Not being

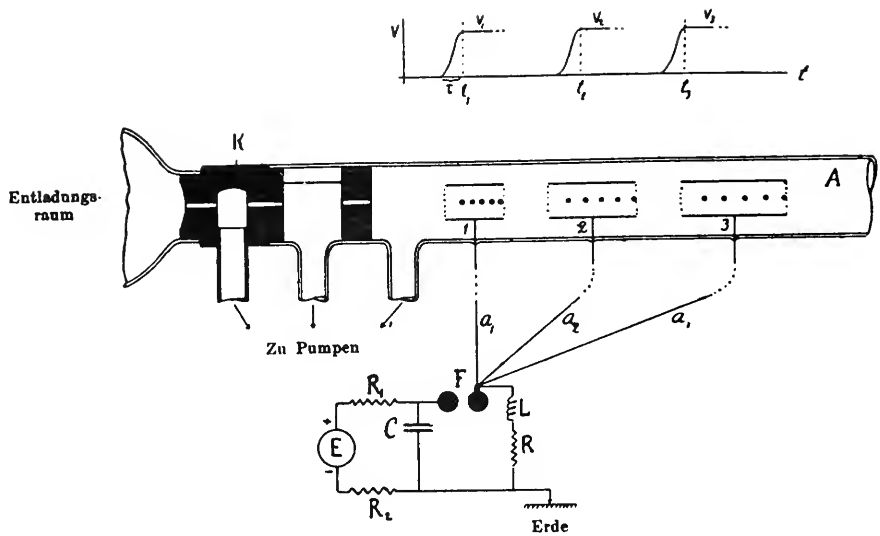


Fig. 1. Diagram of linear accelerator from Professor G. Ising's pioneer publication (1924) of the principle of multiple acceleration of ions.

able to read German easily, I merely looked at the diagrams and photographs of Wideroe's apparatus and from the various figures in the article was able to determine his general approach to the problem — i. e. the multiple acceleration of the positive ions by appropriate application of radio frequency oscillating voltages to a series of cylindrical electrodes in line. This new idea immediately impressed me as the real answer which I had been looking for to the technical problem of accelerating positive ions, and without looking at the article further I then and there made estimates of the general features of a linear accelerator for protons in the energy range above one million volt electrons. Simple calculations showed that the accelerator tube would be some meters in length which at that time seemed rather awkwardly long for laboratory purposes. And accordingly, I asked myself the question, instead of using a large number of cylindrical electrodes in line, might it not be possible to use two electrodes over and over again by bending the positive ions back and forth through the electrodes by some sort of appropriate magnetic field arrangement. Again a little analysis of the problem showed that a uniform magnetic field had just the right properties — that the angular velocity of the ions circulating in the field would be independent of their energy so that they would circulate back and forth between suitable hollow electrodes in resonance with an oscillating electrical field of a certain frequency which now has come to be known as the "cyclotron frequency".

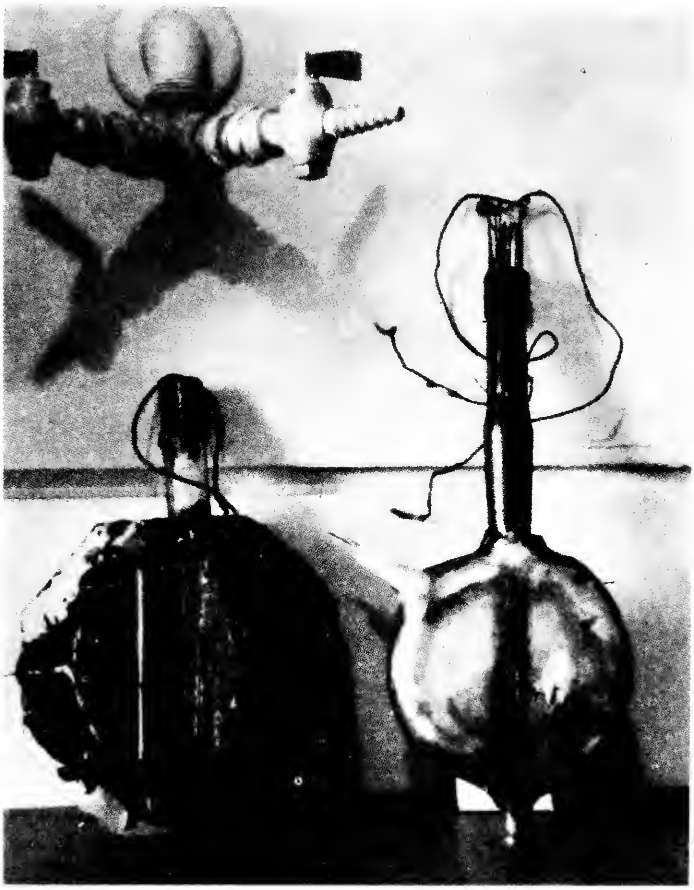


Fig. 2. First crude models of the cyclotron constructed by Edlefsen in 1930.

Now this occasion affords me a felicitous opportunity in some measure to correct an error and an injustice. For at that time I did not carefully read Wideroe's article and note that he had gotten the idea of multiple acceleration of ions from one of your distinguished colleagues, Professor G. Ising, who in 1924 published this important principle. It was only after several years had passed that I became aware of Professor Ising's prime contribution. I should like to take this opportunity to pay tribute to his work for he surely is the father of the developments of the methods of multiple acceleration.

Perhaps you will permit me first of all to show a slide of the diagram of the linear accelerator in his original publication. Fig. 1.

I hope I have not belabored excessively these early incidents of history and now I should like to trace rapidly the evolution of the cyclotron by showing

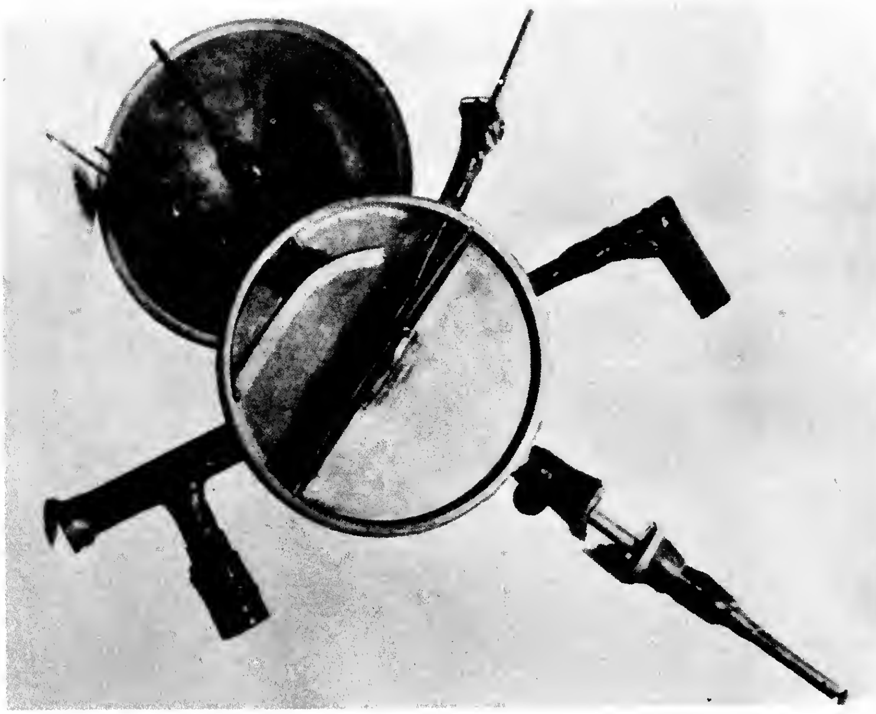


Fig. 3. Working model of cyclotron constructed by M. Stanley Livingston which pointed the way to later developments.

examples of the apparatus in our laboratory as it was developed in the course of time. In doing so, I am afraid I shall not be able to mention all those who deserve great credit for the developments — as from the beginning the work has been a team effort involving many able and devoted co-workers in many laboratories. As I am sure you well appreciate, a great many diverse talents are involved in such developments and whatever measure of success is achieved is dependent on close and effective collaboration.

Although the cyclotron was, so to speak, invented early in 1929, actual experimental work on its development was begun in the spring of 1930 when one of my students, NELS EDLEFSEN, constructed two crude models shown in Fig. 2. One of the models which gave slight evidence of working consisted of two copper duants waxed together on a glass plate with a filament source along the diameter at the center much like later models.

In the fall another student, M. STANLEY LIVINGSTON, continued the development and quickly constructed the model shown in Fig. 3 which, as you see, had all the features of early cyclotrons and which worked very well indeed as

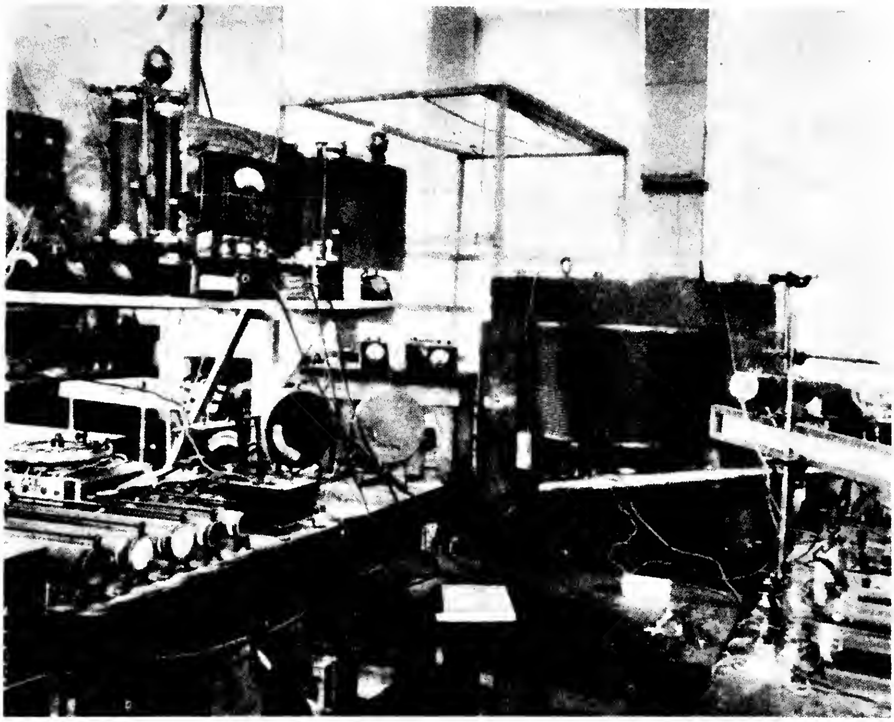


Fig. 4. General view of first cyclotron used in nuclear transformations.

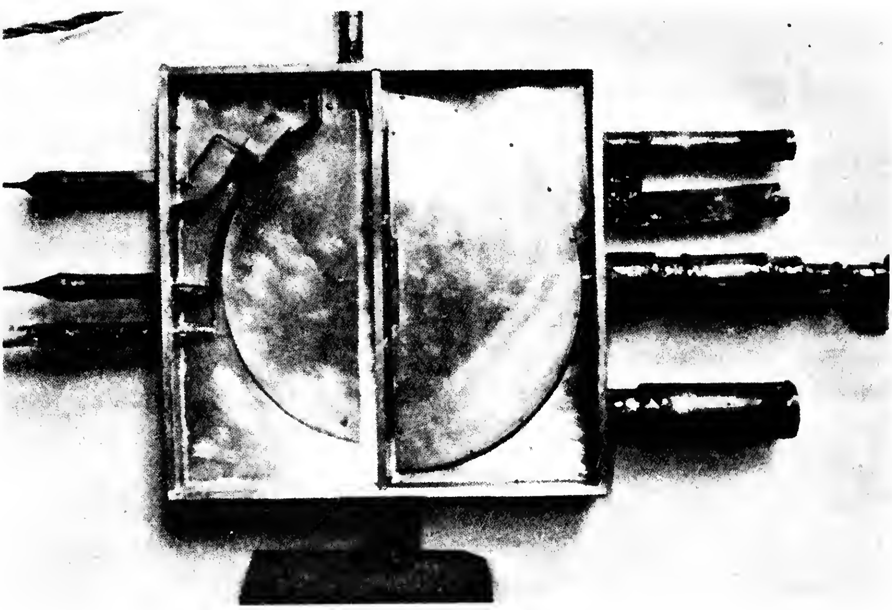


Fig. 5. Vacuum chamber of cyclotron (Fig. 4) which produced 1 million volt protons.

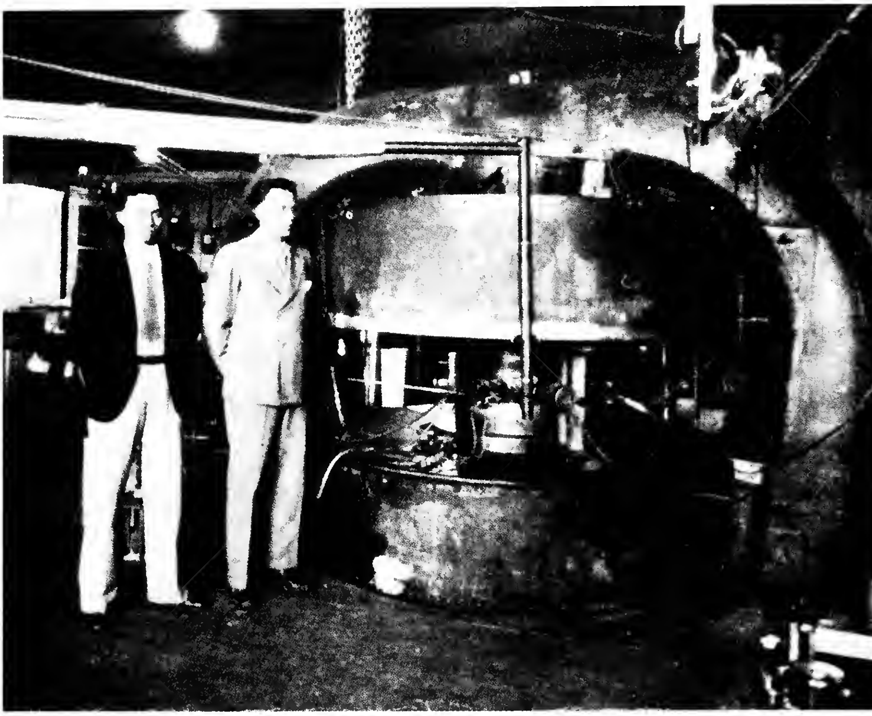


Fig. 6. General view of 27″ cyclotron built by young physicists including M. S. Livingston (left) and E. O. Lawrence (right). The lack of good engineering design is quite evident!

80,000 volt protons were produced with less than 1,000 volts on the semicircular accelerating electrode — now called the "dee".

The next milestone in the development was the construction of a larger model Figs. 4 and 5 which produced protons of the desired energies — in the region of one million electron volts. Livingston and I had the remarkable good fortune of observing that this apparatus was rather more successful than we had expected. For, as you can well imagine, we were concerned about how many of the protons would succeed in spiralling around a great many times without getting lost on the way. We soon recognized that the focussing actions of the electric and magnetic fields were responsible for the relatively large currents of protons that reached the periphery of the apparatus; but we must acknowledge that here again experiment preceded theory!

We were busy with further improvements of the apparatus to produce larger currents at higher voltages when we received word of the discovery by Cockcroft and Walton, which this year has been recognized by the Nobel Prize in physics. We were overjoyed with this news for it constituted definite assur-

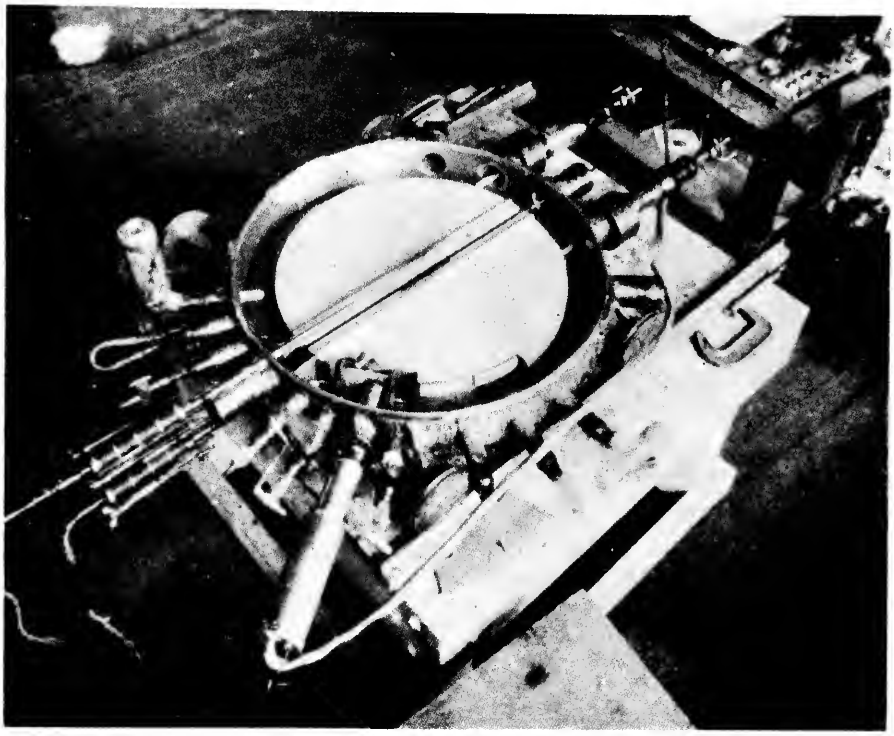


Fig. 7. The chamber of the 27″ cyclotron showing two dees.

ance that the acceleration of charged particles to high speeds was a worthwhile endeavor. As you can imagine, we went ahead with all speed, and it was not long before the disintegration of lithium by protons had been observed with the apparatus.

Now we may proceed rapidly with examples of later developments. Figs. 6 and 7 show the first two dee 27″ cyclotron which produced protons and deuterons of several million volts and was used extensively in early investigations of nuclear reactions involving neutrons and artificial radioactivity.

Again, with this apparatus the discoveries of CHADWICK and the CURIE-JOLIOTS were promptly confirmed. Indeed, looking back it is remarkable that we managed to avoid the discovery of artificial radioactivity prior to their epoch-making announcement: for we tried at first to use Geiger counters in observing nuclear radiations produced by the cyclotron and observed that their background was always variable and large. In those days Geiger counters had the reputation of being unreliable and, rather than looking into the matter of their apparent misbehavior, we turned to ion chambers and linear amplifiers

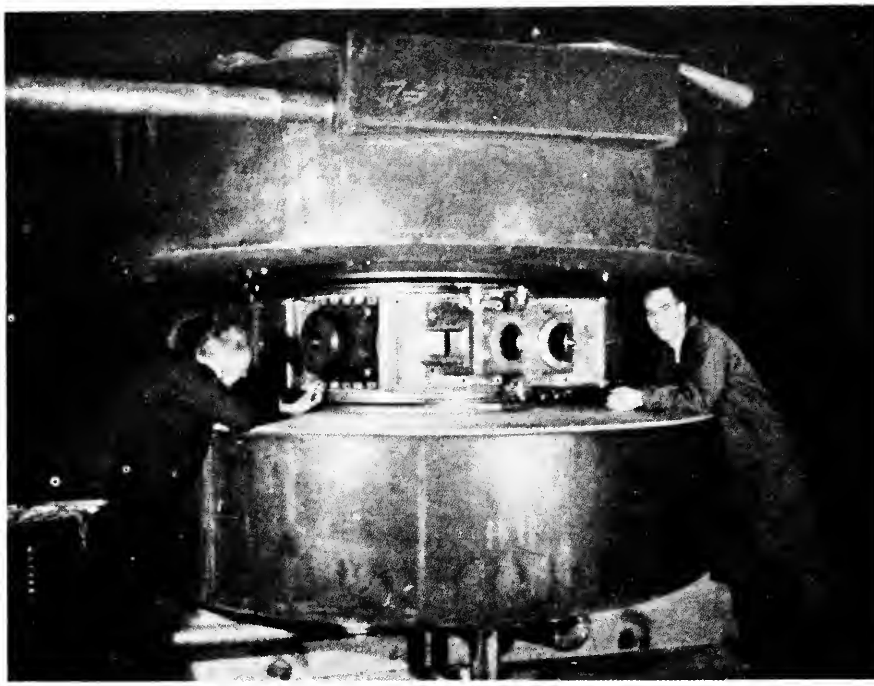


Fig. 8. Early photograph of 60″ cyclotron showing first evidence of good engineering practice introduced into our laboratory by W. M. Brobeck (right) and Donald Cooksey (left).

to observe heavy particle nuclear reactions. Of course, the Geiger counters were simply being faithful to duty and recording the radiations from the artificial radioactive substances and this became immediately apparent after the CURIE-JOLIOT announcement. Again, we were overjoyed at the richness of the domain in the nucleus accessible to particles of several million electron volts energy and there followed a happy period of intensive experimental investigations, which indeed through the years has gained ever-increasing tempo in laboratories the world over.

The next milestone in our laboratory was the construction of the 60″ cyclotron, and this undertaking was greatly strengthened by the joining of our team of WILLIAM BROBECK, a truly outstanding young engineer. Brobeck brought to our laboratory sound engineering practice which from the day he joined us has had a profound effect on developments. To him, more than to any other one individual, goes the credit for the success of the 60″ cyclotron and all subsequent developments. As you can see in Fig. 8, the cyclotron for the first time began to look like a well engineered machine. It was with this machine that the discoveries of the transuranium elements were made which have been

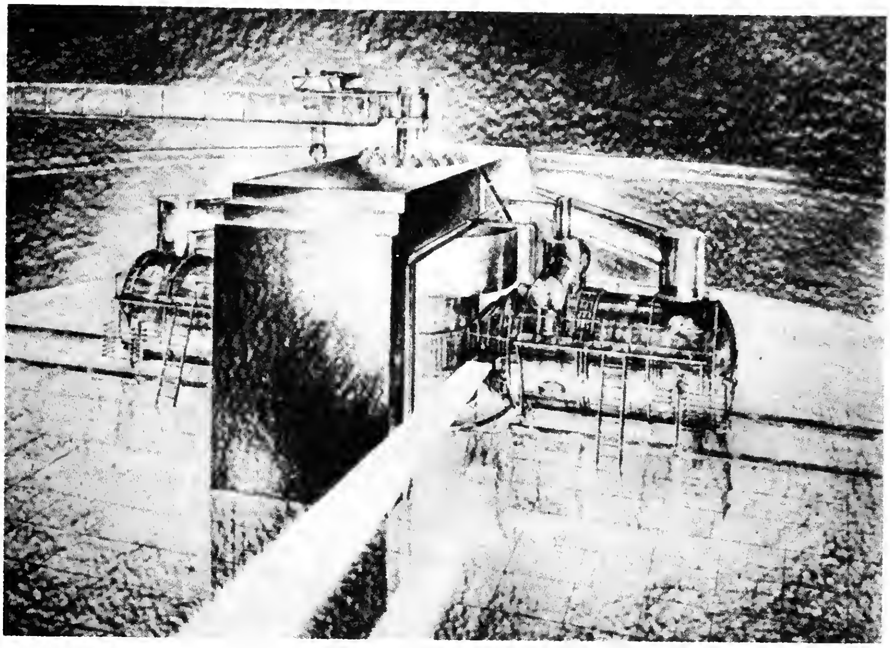


Fig. 9. Artist's sketch of 184″ cyclotron designed by Brobeck before the war to produce 100 million electron volt protons.

rewarded this year by the award of the Nobel Prize in chemistry to McMillan and Seaborg. Perhaps the finest example of a 60″ cyclotron is now in operation at the Nobel Institute here in Stockholm.

Soon our objective was the production of protons and deuterons of much higher energies and Bethe pointed out the difficulty introduced by the relativity increase in mass of the particles as they increase in energy in the course of acceleration which causes them to get out of resonance with an oscillating electric field in a uniform magnetic field.

However, Thomas devised a magnetic field that avoided the limitation discussed by Bethe, and also, of course, it was recognized that one might modulate the frequency in step with the changing angular frequency of the accelerated particles. These two solutions of the technical problem of yet higher energies — the region of 100 million volts — seemed impractical; at least much less practicable than simply so designing the cyclotron that a million volts or more could be applied to the dees, so that the particles would need to circulate around relatively few times in reaching the desired high energies.

Accordingly, just before the war Brobeck and co-workers designed the great 184″ cyclotron shown in Fig. 9.

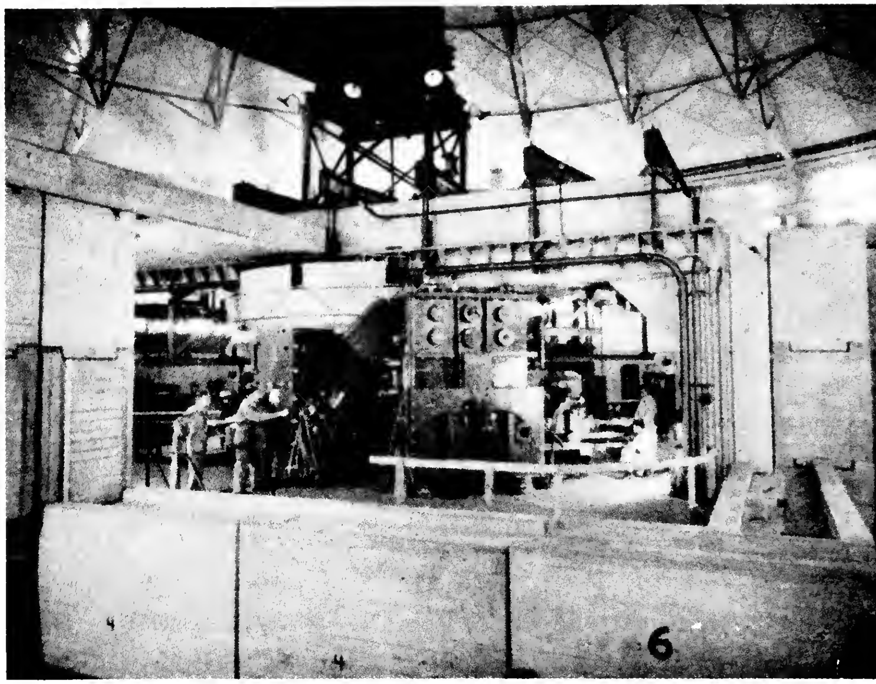


Fig. 10. General view of 184″ synchrocyclotron which produces 340 Mev protons. The concrete shielding, partially removed in this photograph, is 15′ in thickness.

As is well known the war prevented the building of this machine and immediately afterwards McMillan, and Veksler independently a few months earlier, came forward with the principle of phase stability which transformed the conventional cyclotron to a much more powerful instrument for higher energies — the synchrocyclotron. Fig. 10 shows the main features of the Berkeley 184″ synchrocyclotron which produces 340 Mev protons, while there are later and more modern installations, notably at Columbia University and University of Chicago, which produce somewhat higher energies. As I am sure this audience is well aware, a beautifully engineered synchrocyclotron is nearing completion at Upsala.

On completion of the 184″ synchrocyclotron, it was natural that Brobeck should turn his attention to the engineering problem of applying the synchrotron principle to the acceleration of heavy ions, particularly protons, to much higher energies — in the range of billions of electron volts. It was not long before his engineering studies indicated the practicability of producing protons in the energy range well above one billion electron volts.

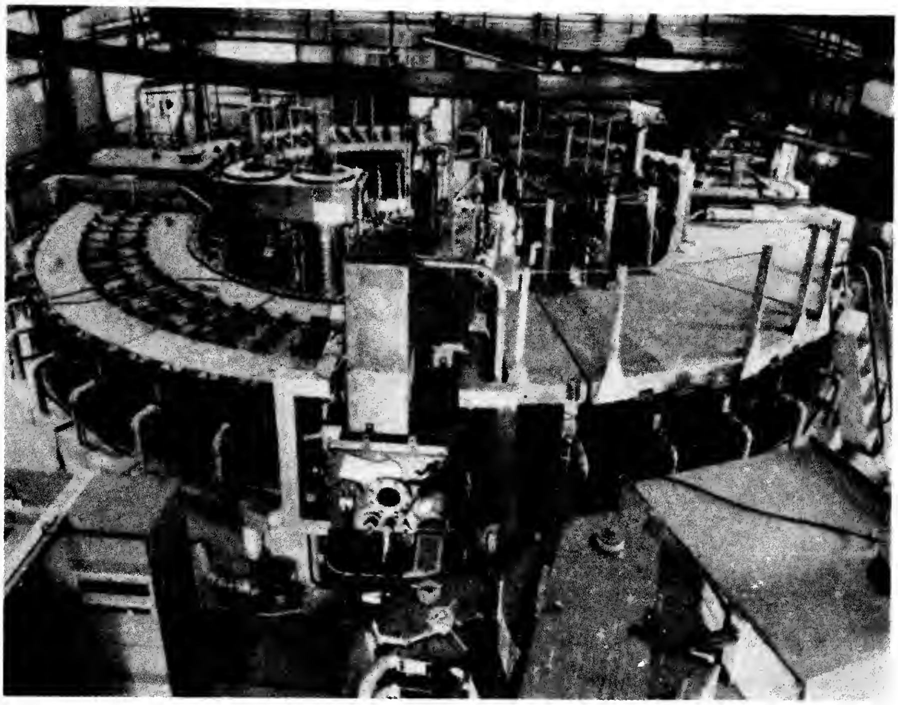


Fig. 11. One-quarter scale operating model of 6 Bev proton synchrotron.

With the extensive developments in the atomic energy field, large funds became available for research purposes — much larger than seemed possible before the war — and indeed, as soon as all concerned were convinced of the practicality of building a proton synchrotron for several billion electron volts, the construction of two installations was begun, one at Brookhaven for about 3 billion electron volts and a second at Berkeley for about twice this energy.

The first step in these large undertakings was to build a substantial operating model to test out the theory of the proton synchrotron, as well as the engineering principles of design. Accordingly, a quarter scale operating model was constructed and is shown in Fig. 11. A small cyclotron was designed to produce large current pulses of 1 Mev protons which were injected into the "race track" of the synchrotron by an appropriate magnetic and electrostatic deflecting system which can be seen in the foreground of Fig. 11. This model worked as expected and provided a great deal of practical data giving confidence that the full scale machines will function successfully and satisfactorily.

It is hardly appropriate here to describe either the Brookhaven or Berkeley proton synchrotrons (the former is called the cosmotron and the latter is called

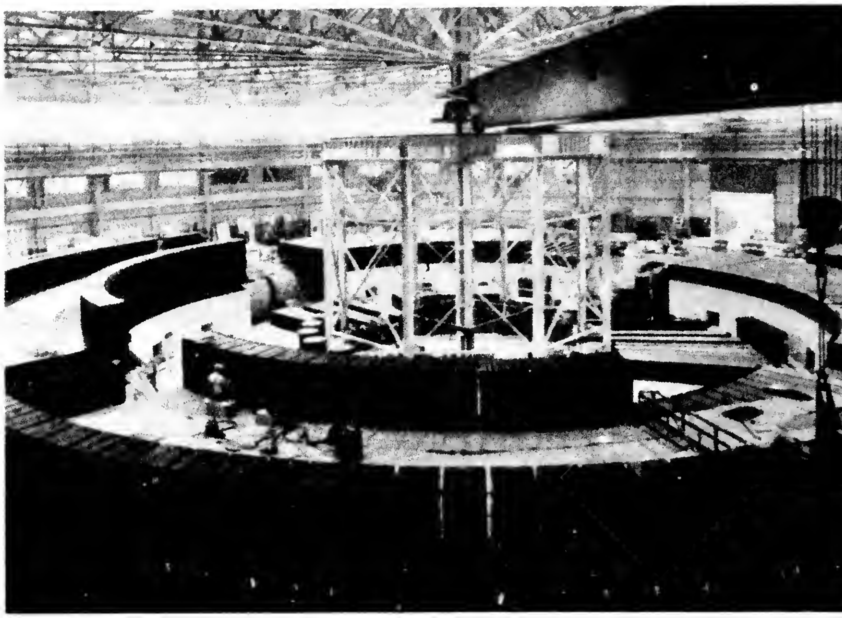


Fig. 12. General view of "race track" magnet in process of assembly for 6.3 Bev proton synchrotron or "bevatron".

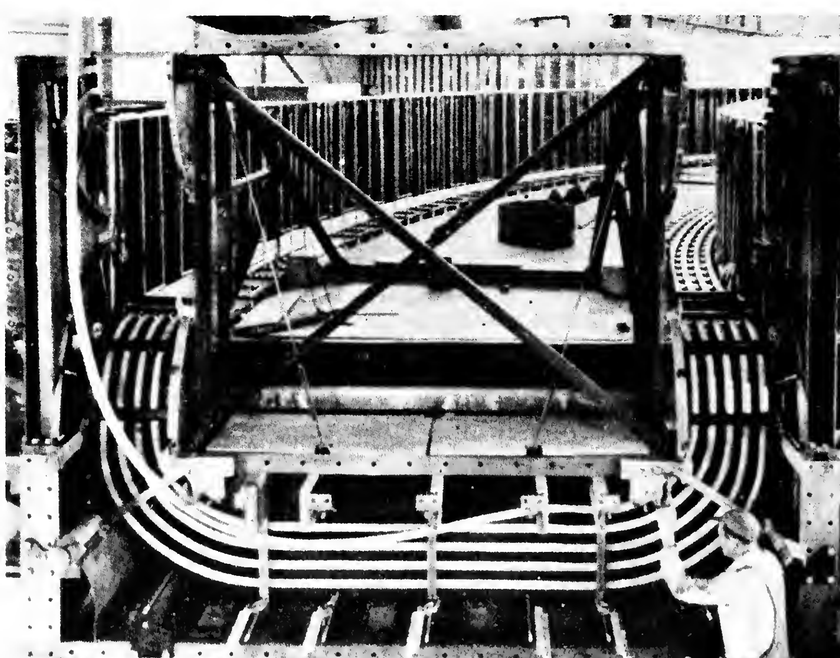


Fig. 13. Showing coil winding of bevatron magnet.

Fig. 14. The size of the bevatron magnet is here indicated. Left to right (E. O. Lawrence, W. M. Brobeck, H. A. Fidler and D. Cooksey).

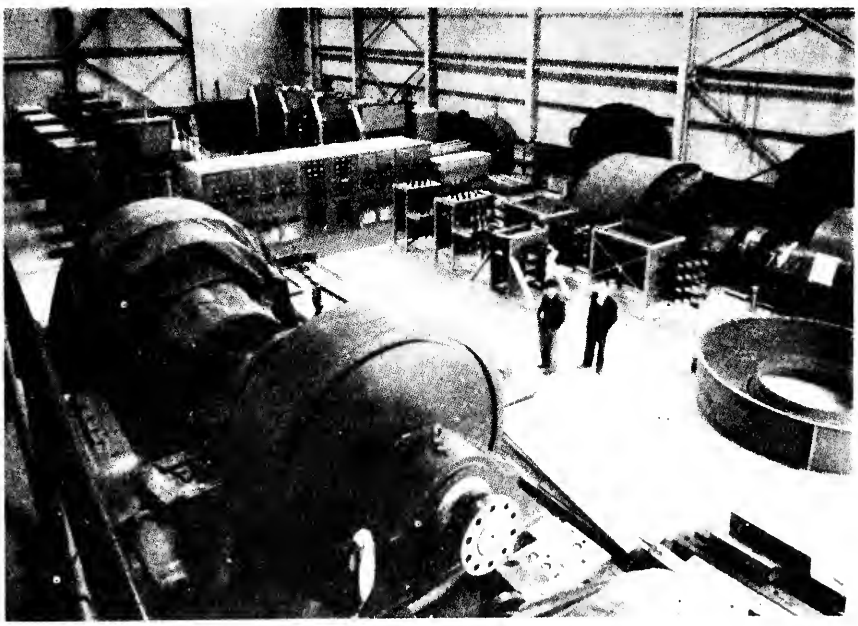


Fig. 15. Bevatron motor generator equipment.

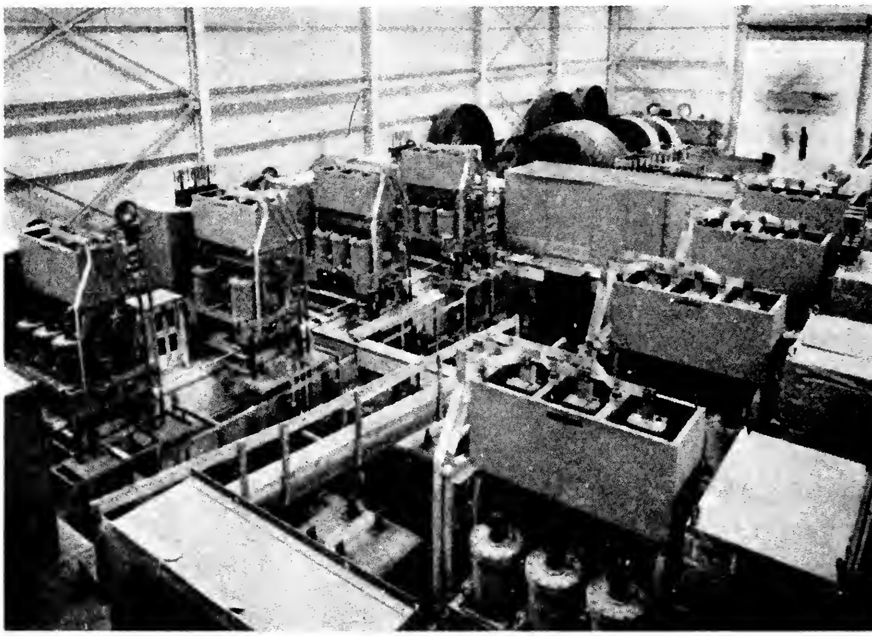


Fig. 16. Ignitrons and associated switchgear for bevatron motor generator.

the bevatron) but perhaps it is of interest to show a number of photographs which display the general features of this great machine. Figs. 12, 13, 14, 15 and 16.

Now that we shall soon have 5 or 10 Bev particles in the laboratory, what possibilities are there for going on higher to 50 or 100 Bev? One answer is that the limitation of the bevatron is largely a financial one. With a correspondingly larger expenditure higher energies surely can be reached.

But I should like to close by emphasizing that a more feasible, if not more interesting, approach to the problem of higher energy nuclear projectiles is the acceleration of multiply charged heavier ions such as C^{6+}, or Ne^{10+}. Already extraordinarily interesting nuclear reactions have been produced by the acceleration of C^{6+} ions to 120 Mev in the 60″ cyclotron and such particles in the Berkeley bevatron would be accelerated to more than 36 Bev. Since in the cosmic radiation such heavy particles play an important rôle, they will surely be produced in the bevatron some day, contributing to further progress in our understanding of nature.

These "machines" are used for two purposes: to "see" fundamental particles of matter, and to produce new ones. Though published in 1958, this article is still an excellent introduction to the basic design used to build many current accelerators.

8 Particle Accelerators

Robert R. Wilson

Article in Scientific American, 1958.

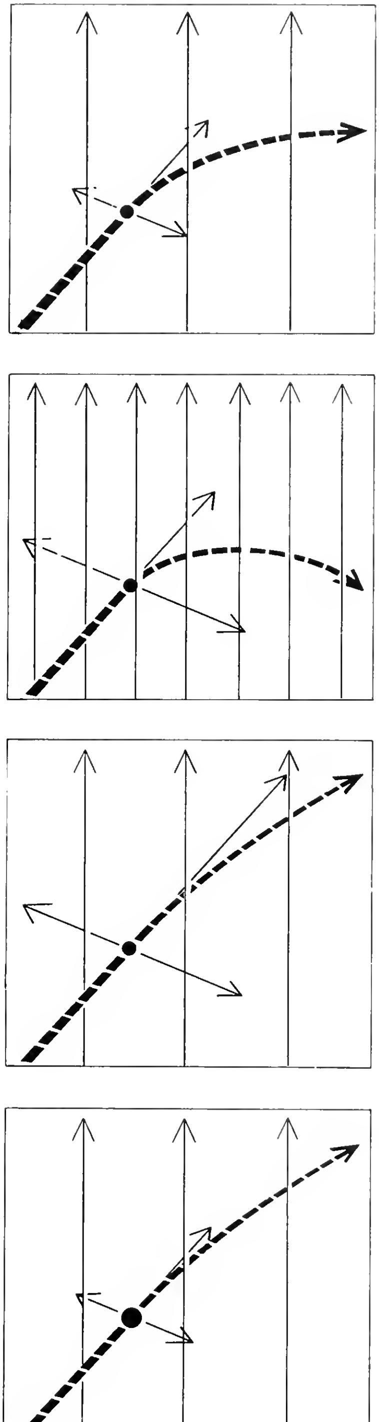
From time to time in the course of history men have been swept up by intense currents of creative activity. In the pyramids of Egypt, in Greek sculpture and in Florentine painting we find monuments to such bursts of expression. My favorite example is the Gothic cathedrals that so magically sprang up in 12th- and 13th-century France, for I like to relate that magnificent preoccupation with construction to an obsession of our own time—the building of nuclear accelerators.

Like nuclear physics today, religion at that time was an intense intellectual activity. It seems to me that the designer of an accelerator is moved by much the same spirit which motivated the designer of a cathedral. The esthetic appeal of both structures is primarily technological. In the Gothic cathedral the appeal is primarily in the functionality of the ogival construction—the thrust and counterthrust that is so vividly evident. So, too, in the accelerator we feel a technological esthetic—the spirality of the orbits of the particles, the balance of electrical and mechanical motion, the upward surge of forces and events until an ultimate of height is reached, this time in the energy of the particles. In both cases we find the architects working at the very limit of technical knowledge. In both there is intense competition between localities, regional and national. Both structures are expensive: a really large accelerator can cost $100 million; the cost of a cathedral, in terms of medieval economics, was possibly higher.

But where a cathedral was a community enterprise, with many people in the region participating in its financing and construction, and nearly everyone in its enjoyment, an accelerator is esoteric. Its presence in a community is usually unknown and unsung. Few are the workers who help to build it, and fewer still are those who use it.

So the accelerator building boom goes on largely unnoticed, but at a quickening pace. Cyclotrons, the original "atom smashers," are now dotted almost all over the globe. They have evolved into synchro-cyclotrons, and have reached their culmination in three giant machines, one at the University of California in Berkeley, another at the European Organization for Nuclear Research (CERN) in Switzerland and another in the U.S.S.R. These machines accelerate

PROTON SYNCHROTRON IN GENEVA is designed to yield 25 bev. Shown here is a section of the interior of its ring building. This structure is approximately 660 feet in diameter.



MAGNETIC FORCE on moving charged particles (*black dots*) is indicated by arrows pointing down and to right. Upward arrows show the speed of the particles and colored arrows the direction of the field. Large dot at the bottom represents a heavier particle.

protons to energies of between 600 and 700 million electron volts (mev). Synchrotrons, another development, are even bigger and more powerful. The Cosmotron, a 2,200-ton monster at Brookhaven National Laboratory which emits 3-billion-electron-volt (bev) protons, is small compared to the 6-bev, 10,000-ton Bevatron at Berkeley. This in turn is topped by the 10-bev, 36,000-ton Phasotron in the U.S.S.R. Two even larger machines are under construction at Brookhaven and CERN; they are designed to produce protons of 25 to 30 bev. And still bigger accelerators are being planned.

Nuclear Microscopes

Why? What is the purpose behind this almost feverish effort to build more and bigger machines? Perhaps the simplest answer is that accelerators are the microscopes of nuclear physics. We usually think of an accelerator as a sort of gun, producing high-speed particles which bombard the nucleus of the atom. But since particles are known to have wave properties, it is equally appropriate to say that the accelerator shines "light" on the nuclei, enabling us to "see" them.

Now the resolving power of a microscope, *i.e.*, its ability to distinguish small objects, depends on the wavelength of the light it employs. The shortest wavelength of visible light is about four 100,000ths (4×10^{-5}) of a centimeter; with these waves one can perceive a microbe of about the same length.

To examine smaller things, biologists now use the electron microscope. The wavelength of a particle depends on its mass and its energy. At a few thousand electron volts—the energy at which electron microscopes operate—an electron has a wavelength some 10,000 times shorter than that of visible light (about 10^{-9} centimeter). With these waves one can begin to see the details of molecules.

The nucleus of an atom is about 10^{-12} centimeter in diameter. This is the wavelength of a proton with an energy of 1 mev. To "see" the nucleus we therefore need a 1-mev proton "microscope," and to make out some of its internal details we need some 10 to 20 times as much energy. Thus a laboratory interested in classical nuclear physics will invariably have a Van de Graaff accelerator or a cyclotron operating in the range of 1 to 20 mev.

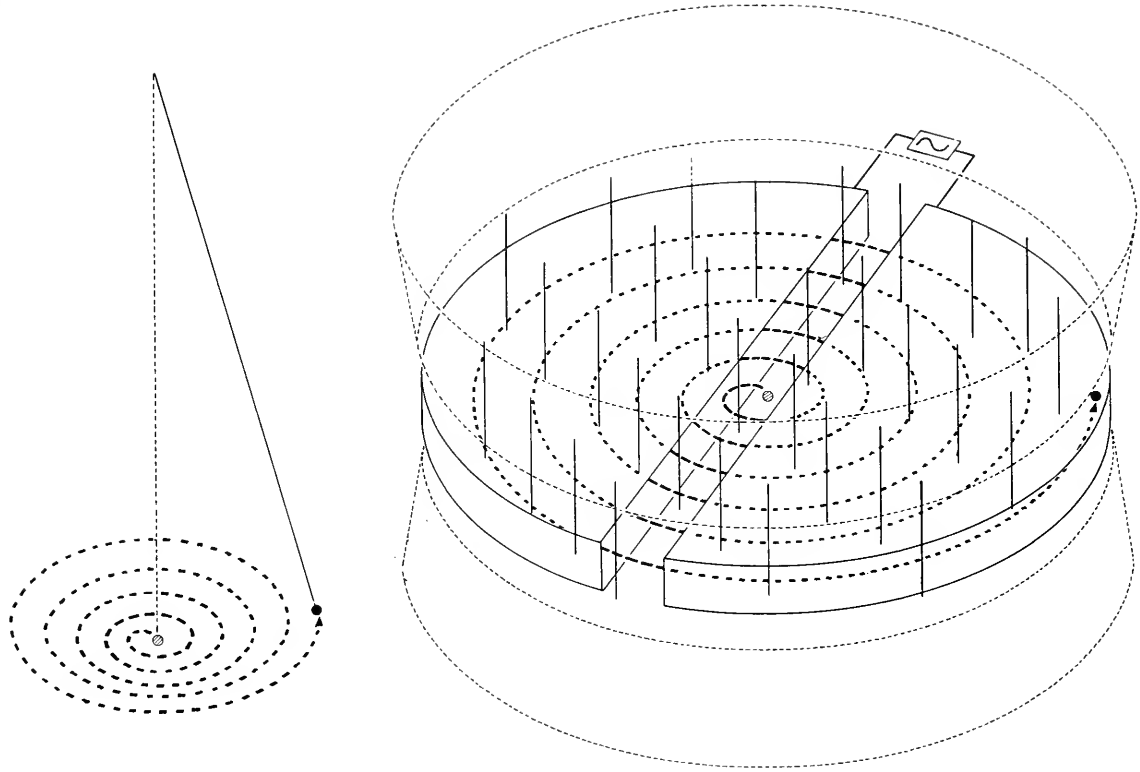
But physics has pushed beyond this point. At present many of us are interested not in the nucleus as a whole but in the structure of the protons and neutrons (nucleons) of which it is composed. It is the old problem of worlds within worlds, for the proton itself turns out to have a rich structure. It is perhaps 10^{-13} centimeter in diameter, and to resolve it requires an energy of several hundred mev. To see it in as fine detail as we can see the structure of the nucleus we must have still higher energy. It is for this reason that the 25- to 30-bev machines are under construction. If and when the structure of the proton is known, will its component parts turn out to have their own structure? Very possibly so, and if they do, machines of higher energy will be built to explore that structure.

The microscope analogy does not tell the whole story. When we get to sufficiently short wavelengths (*i.e.*, when the bombarding particles in our accelerators reach sufficiently high energy), we not only see particles, but we also make new ones. These new particles are created out of energy. At 1 mev an electron has enough energy to create a pair of particles—an electron and a positron. At 150 mev it makes pi mesons (pions) when it collides with a nucleon. Our 1-bev electron accelerator at Cornell University produces more massive particles: K and lambda mesons. The Bevatron, which produces 6-bev protons, is able to create antiprotons, antineutrons and still heavier particles such as xi and sigma mesons.

Thus as the energy of the machines has increased it has become possible to create more and heavier new particles. Obviously the exciting next step is to attain even higher energies, and then to see what sort of monster particles are created. One has the very strong feeling that new particles will indeed show up. It may well turn out that they will prove to be only complexes of particles which we already understand; however, it is exactly to answer such questions that we are building the machines.

Originally we constructed our accelerators in order to search for the ultimate in elementary particles. We expected these particles to be fragments and hence to be successively smaller; it was to improve our definition of them that we went to higher energies. Ironically the fragments now seem to get larger. One has the uneasy feeling that new machines make new particles which lead to the construction of new machines, and so on *ad infinitum*. In fact, there may be lurking here a new kind of indeterminacy principle which will inherently limit our knowledge of the very small.

So much for the reasons why accelera-



CYCLOTRON'S OPERATION is like that of a circular pendulum (*left*) in which the weight is pushed repeatedly to give an ever-widening swing. The schematic diagram at the right shows a particle (*dot*) spiraling within two D-shaped electrodes. The magnetic pole pieces which provide the guiding field (*colored lines*) are outlined in light broken lines. The particles are accelerated by an oscillating electric field between the dees. The generator which produces the field is shown as a wavy line within a rectangle (*top*).

tors are built. Let us turn to the machines themselves. All of them operate on the same fundamental principle: charged particles (electrons or positive ions usually protons) are put into an electric field which exerts a force on them, pushing them to high speeds and energies. (The electron volt, in which the energy is usually measured, is the energy acquired by a particle with one electronic unit of charge accelerated by a potential difference of one volt.) The simplest form of accelerator is a pipe along which a steady electric field accelerates the particles. This is the well-known Van de Graaff machine. To obtain higher energies a long pipe may be used with several accelerating electrodes which kick the particles to higher and higher speeds as they travel down the tube [see "The Linear Accelerator," by Wolfgang Panofsky; SCIENTIFIC AMERICAN, October, 1954]. But to attain a really high energy by this method would require an extremely long pipe. To get around this difficulty the particles can be made to travel in a circular or spiral path which brings them back through the same electrodes where the accelerating voltage is applied again and again.

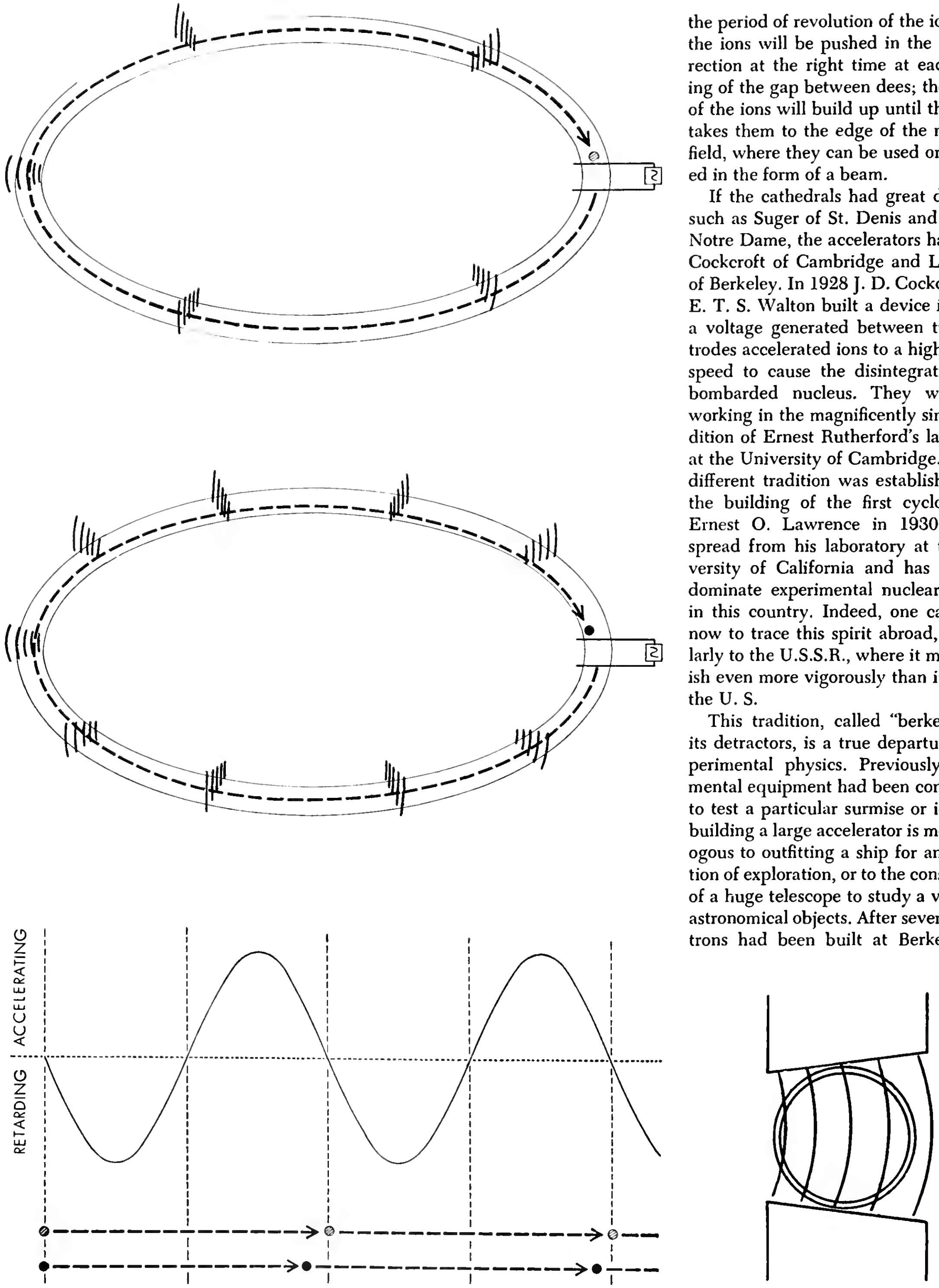
It is with such circular machines that we are chiefly concerned in this article. In these machines the circular motion is brought about by magnetic fields. A magnetic field exerts a force on all electric charges that move through it; the force is always at right angles to the direction of the charges' travel. It is the same kind of force that acts on a stone whirled at the end of a string. The magnetic field, like the string, forces the particles to move in a circular path. The stronger the field, the sharper the curvature of the path; on the other hand, the faster or heavier the particle, the less it is curved by a given field [*see diagrams on opposite page*].

The simplest and oldest type of accelerator to make use of magnetic bending is the cyclotron. The operation of this machine can be most easily visualized by imagining a weight suspended by a string and pushed so as to describe a circular motion. As with any pendulum the time required to complete a full circular swing is the same whether the circle is small or large. Thus if the weight is pushed rhythmically it will move outward in an ever-widening circle, returning to the pushing point in the same time on each revolution [*see diagram above*]. So it is in the cyclotron: each ion whirls inside of two semicircular electrodes or "dees," getting an electrical push when it passes from one to the other. A vertical magnetic field provides a constant inward push and, like the string, holds the ion in a circular path and guides it back to the gap between the dees, where it is given another electrical push. The velocity of the ion then becomes greater and, as a result of its inertia, the curvature of the circular path caused by the magnetic field becomes larger. The time taken to traverse a full circle is the same no matter how big the radius, because the increase in speed just compensates for the increase in path-length per turn. Now if the voltage across the dees is made to oscillate rapidly, and if its period is adjusted so that it exactly matches

the period of revolution of the ions, then the ions will be pushed in the right direction at the right time at each crossing of the gap between dees; the energy of the ions will build up until their path takes them to the edge of the magnetic field, where they can be used or extracted in the form of a beam.

If the cathedrals had great designers such as Suger of St. Denis and Sully of Notre Dame, the accelerators have their Cockcroft of Cambridge and Lawrence of Berkeley. In 1928 J. D. Cockcroft and E. T. S. Walton built a device in which a voltage generated between two electrodes accelerated ions to a high enough speed to cause the disintegration of a bombarded nucleus. They were still working in the magnificently simple tradition of Ernest Rutherford's laboratory at the University of Cambridge. A quite different tradition was established with the building of the first cyclotron by Ernest O. Lawrence in 1930. It has spread from his laboratory at the University of California and has come to dominate experimental nuclear physics in this country. Indeed, one can begin now to trace this spirit abroad, particularly to the U.S.S.R., where it may flourish even more vigorously than it does in the U. S.

This tradition, called "berkelitis" by its detractors, is a true departure in experimental physics. Previously experimental equipment had been constructed to test a particular surmise or idea. But building a large accelerator is more analogous to outfitting a ship for an expedition of exploration, or to the construction of a huge telescope to study a variety of astronomical objects. After several cyclotrons had been built at Berkeley, the



SYNCHROTRON restricts particles to a nearly circular path by means of a magnetic field (*colored lines*) which grows stronger as the particle energy increases. At top an electron (*hatched circle*) is in an orbit that brings it to the accelerating gap (*right*) just as the voltage changes from accelerating to retarding (*curve at bottom*). In the center drawing the field is made stronger and the electron (*black circle*) is bent more strongly, following a shorter path and arriving at the gap in time to get a push. After a number of pushes it spirals out to the original path. The cross section at bottom right shows magnetic pole pieces around the doughnut.

students and associates of Lawrence traveled far and wide to spread the gospel. By World War II they had helped to build cyclotrons not only at universities in the U. S., but also in several other countries. The biggest of these machines produced protons of about 10 mev. As we have seen, this is an appropriate energy for exploring the nucleus as a whole, but not for examining its parts. Just before the war Lawrence had begun to build a giant cyclotron, to enter the energy region above 100 mev, with which he could start to probe nucleons.

The Synchrotron

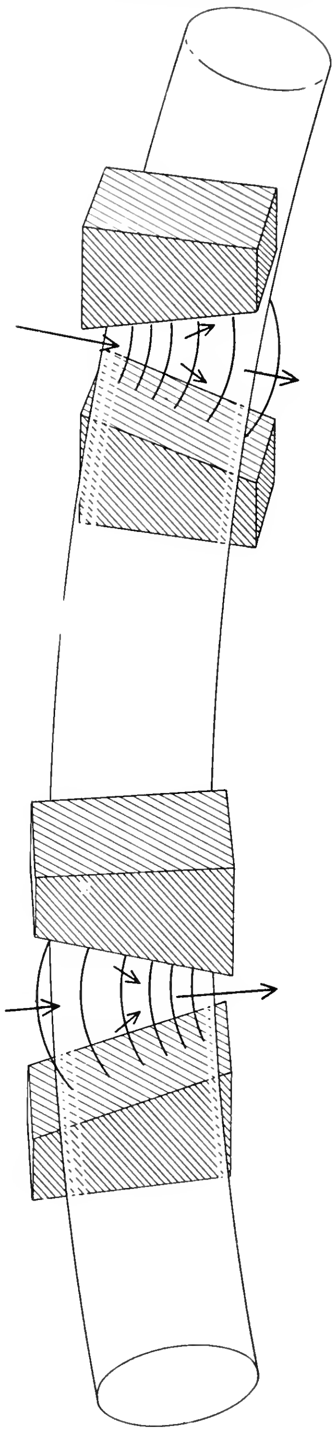
It was characteristic of Lawrence that he went ahead despite a prevalent conviction that the energy limit of the cyclotron was about 20 mev. This conviction was based on an effect predicted by Albert Einstein's theory of relativity: particles traveling at nearly the speed of light will increase in mass. At 20 mev a proton has entered this "relativistic" region, and further increases in energy will result not so much in greater speed as in greater mass. When this happens, the particle in a cyclotron begins to fall behind schedule as it spirals farther outward, and it no longer arrives between the dees at the right time to get a push from the oscillating voltage.

The war interrupted work on Lawrence's big machine. Its huge magnet was used to separate isotopes of uranium for the atomic-bomb program. At the end of the war V. I. Veksler of the U.S.S.R. and E. M. McMillan of the University of California independently and almost simultaneously enunciated the so-called synchrotron principle. This principle showed the way to accelerating particles into the completely relativistic region. It was exactly the sort of *deus ex machina* that Lawrence had envisioned when he gambled some $1 million in starting his big cyclotron. The principle was immediately adopted. A successful synchro-cyclotron was built which produced protons in the region of 100 mev (eventually 730 mev). In the next few months a number of important features of the proton were discovered.

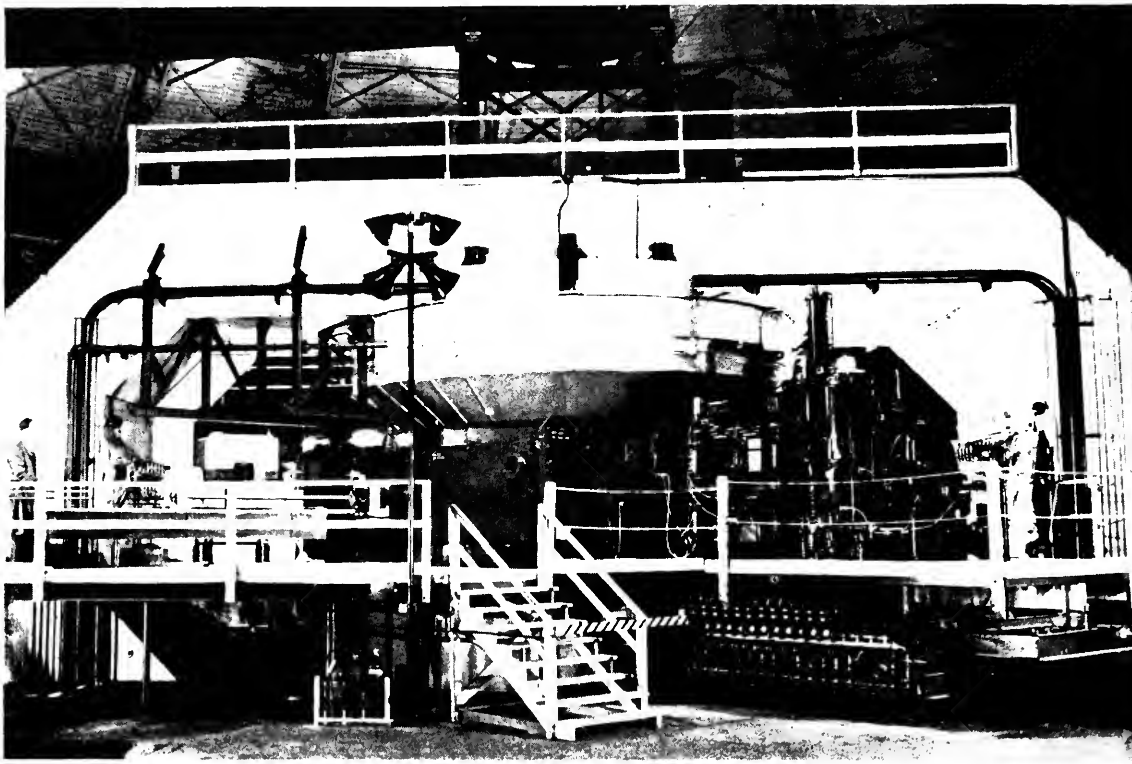
To understand the synchrotron principle, it is easier to consider its application in the electron synchrotron rather than in the more complicated synchro-cyclotron. Some half-dozen of these electron accelerators, with maximum energies of about 300 mev, were also built just after the war.

In a synchrotron electrons travel on a circular orbit inside a narrow doughnut-shaped vacuum vessel. At one point in the doughnut is a pair of accelerating electrodes across which there is an oscillating voltage like that in the cyclotron. A ring-shaped magnet surrounding the doughnut produces a field which forces the particle to travel on orbits close to the center of the tube [*see diagram on opposite page*]. The electrons are injected into the doughnut from a small linear accelerator at an energy of about 2 mev. At this energy their speed is some 98 per cent of the speed of light; hence they cannot travel much faster. To make matters simpler let us assume that the speed is exactly the speed of light and that the whole increase in energy goes into mass. Now imagine an electron in a circular orbit at the center of the doughnut. The electron is held there by a constant magnetic field. Also imagine that our oscillating voltage is applied, but that the electron crosses the accelerating gap just at the time when the voltage falls through its zero value. The frequency of the voltage is made the same as that of the electron traveling around its orbit at the constant speed of light. The electron now passes the gap on all subsequent turns just as the voltage becomes zero. Thus nothing happens; the electron remains on its orbit and keeps the same energy. Now we increase the magnetic field slightly. Since the energy (mass) is still the same, the particle is forced into a sharper curve, *i.e.*, its orbit gets smaller. But because the orbit is smaller and the speed is constant, the time it takes the electron to return to the accelerating gap is shorter. Hence the electron arrives slightly before the voltage has fallen to zero; it is accelerated slightly. On the next turn, if the energy is still not large enough, the orbit will still be too small: the electron will arrive still earlier and be accelerated even more. Eventually the energy will increase enough (that is, the electron will get heavy enough) so that it is bent less sharply and edges out to its original orbit. If the energy should become too great, the orbit will be too big and the time it takes the electron to make each turn will be too long. This will cause the electron to drop behind the accelerating voltage and be pushed backward so that it will lose energy. Thus we have a beautiful automatic device for keeping the electron on the right orbit, or at least oscillating around the right orbit. That is all there is to the synchrotron principle or, as it is sometimes called, phase focusing.

Now we can see that, if the magnetic field of the synchrotron is increased continuously, the energy of the electrons



STRONG FOCUSING is produced by magnetic fields which are alternately bowed out and in. Horizontal arrows show radial forces on the particles at inner and outer edges of the field. Slanted arrows represent forces which focus or defocus particles vertically.



SYNCHRO-CYCLOTRON at the Berkeley Radiation Laboratory of the University of California is now the most powerful machine of its kind. A modification of its design last year increased the energy of its proton beam to 730 million electron volts (mev).

ELECTRON SYNCHROTRON was photographed in the author's laboratory at Cornell University while its guiding magnet was under construction. Machine, which produces an energy of 1 bev, is the first to use strong focusing. Accelerating electrodes are at right.

will also increase continuously; the electrons will receive energy at just the right rate to keep them on the central, or synchronous, orbit. In practice electrons can be injected into the doughnut when the magnetic field is rather weak (about 10 gauss) and ejected when the field is quite strong (more than 10,000 gauss). A synchrotron with a large enough radius can accelerate electrons up to energies of about 10 bev. There are now about six machines, built or being built, which are designed to yield electron energies between 1 and 1.5 bev. At Cambridge, Mass., a 6-bev electron synchrotron is being constructed by a joint Harvard University–Massachusetts Institute of Technology group.

Let us return to the synchro-cyclotron. It works in essentially the same way as a synchrotron but it is shaped like a cyclotron. Instead of a varying magnetic field it has a constant field, but the frequency of the accelerating voltage applied to the dees is varied. This means that the synchronous orbit of the protons is not a fixed circle but a growing spiral.

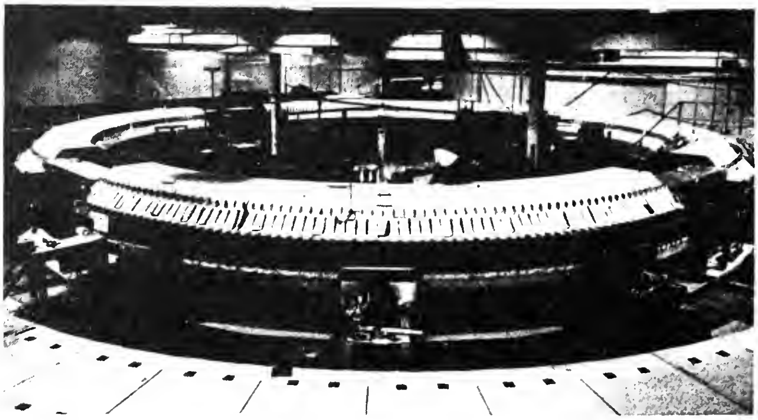
In another class of accelerators, the proton synchrotrons, both the magnetic field and the frequency of the accelerating voltage are varied. The increasing field counteracts the protons' tendency to spiral outward as they get up to relativistic energies, and the orbit is again a fixed circle. Above about 5 bev the protons are traveling practically at the speed of light, and from here on the proton synchrotron works just like an electron synchrotron.

If I may extend the figure of speech with which I began this article, each kind of accelerator has its own architectural style. To me synchro-cyclotrons are baroque. Proton synchrotrons are definitely Romanesque, although their rounded arches are horizontal. Electron synchrotrons have a lightness and grace that could only be Gothic.

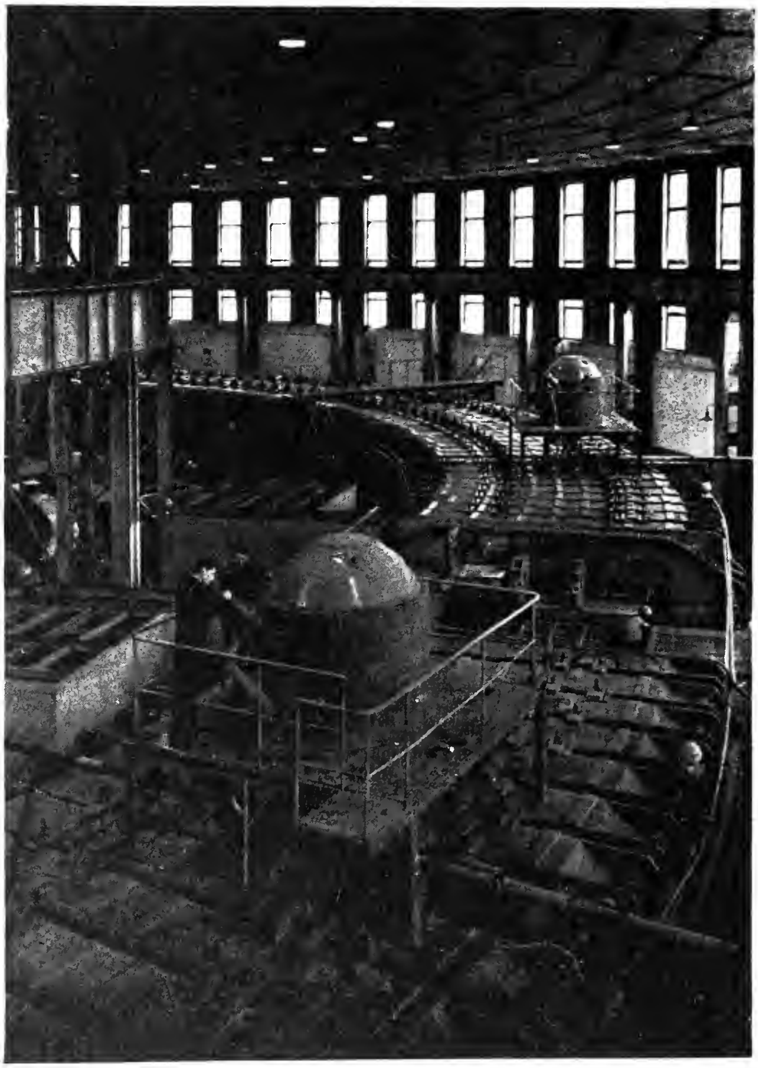
The Newer Machines

This brings us more or less up to date in the evolution of accelerators. We may now ask whether we are near the end of this movement in physics or still at its beginning. The field still has tremendous vigor, and it is my guess that we are at about the same stage as the cathedral builders were after they had completed Notre Dame of Paris. The significant innovations were behind them, but most of their masterpieces were yet to come.

Early in this article I mentioned that two machines now under construction, one at Brookhaven National Laboratory



COSMOTRON, the 3-bev proton synchrotron at Brookhaven National Laboratory, was the first one of the multi-bev accelerators. Its 2,200-ton magnet has an inside diameter of 60 feet.



PHASOTRON is a 10-bev proton synchrotron in the U.S.S.R. Its magnet, of which a portion appears in this photograph, weighs 36,000 tons and is approximately 200 feet in diameter.

FFAG (fixed-field alternating-gradient) design is embodied in an electron accelerator built as a model for a larger proton machine at the laboratory of the Midwestern Universities Research Association in Madison, Wis. The dark spiral sectors are the magnets.

and the other at CERN in Geneva, will produce protons of 25 to 30 bev. Both of these machines are proton synchrotrons; each will cost between $20 million and $30 million. The diameter of the orbit traveled by their protons will be nearly 1,000 feet!

These machines were made possible by the discovery at Brookhaven of a new principle called strong focusing [see "A 100-Billion-Volt Accelerator," by Ernest D. Courant; SCIENTIFIC AMERICAN, May, 1953]. This principle involves a reshaping of the guiding magnetic field so that the particles are held much closer to their ideal orbit. It means that the doughnut can be thinner, and the surrounding magnet smaller and lighter.

Until now we have considered only the radius of the orbit, *i.e.*, the size of the circle on which the particles travel. However, the particles can not only drift in and out but also up and down; thus they must be focused vertically as well as horizontally. In old-style synchrotrons the lines of force in the magnetic field are bowed slightly outward [*see diagram on page 6*]. This has the effect of forcing particles back toward the center line when they move above or below it. But the bowed field gets somewhat weaker with the distance from the center line. Hence a particle that wanders too far from the center line is not strongly pushed back toward it.

In strong focusing the field is broken into sectors which are alternately bowed outward and inward [*see diagram on page 7*]. The sectors bowed outward provide sharp vertical focusing, but are even worse than the old field-shape at bringing a particle in from an orbit that is too large. In other words, they do not focus radially. On the other hand, the sectors bowed inward increase in strength as the radius gets bigger, and provide strong radial focusing. Vertically, however, they have the wrong effect on the particles, tending to spread rather than to focus them. It turns out that each of the defocusing influences is overbalanced by the focusing effect of the other sector; the net result is a much more tightly restricted beam. This method of focusing was successfully used in the Cornell 1-bev electron synchrotron, and it will be applied in the 6-bev Harvard-M.I.T. electron synchrotron.

Not to be outdone by CERN and Brookhaven, the U.S.S.R. has announced that it will build a 50-bev strong-focusing proton synchrotron. The magnet will weigh about 22,000 tons and will have a diameter of 1,500 feet. It would seem that whatever we do, our Soviet friends can do too—and with a factor of two in their favor.

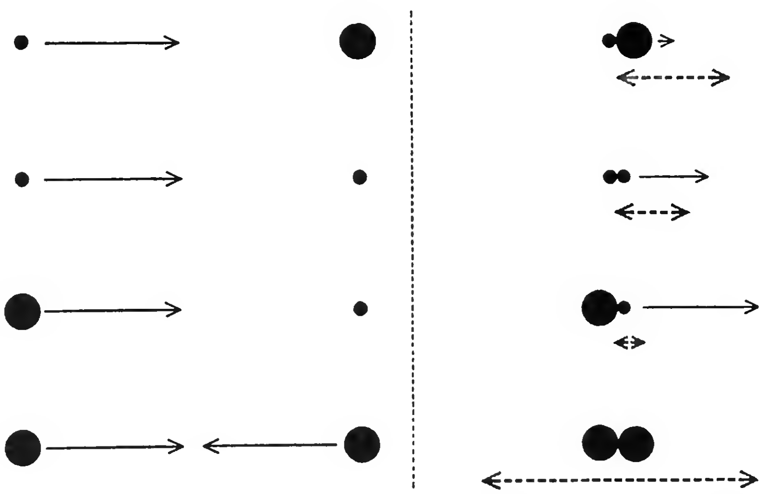
"FFAG"

The most exciting recent development in this country has been the "fixed-field alternating-gradient" accelerator proposed by Keith R. Symon of the Midwestern Universities Research Association (MURA). The so-called FFAG machine is really a rococo cyclotron in which the magnetic field is shaped in such a way as to allow the cyclotron to work into the high-energy relativistic region. We have already seen how the ordinary cyclotron is limited to accelerating protons to about 20 mev. When this limitation was first pointed out in 1938, L. H. Thomas of the Ohio State University suggested a way to get around it. He proposed to scallop the pole tips of the cyclotron magnet so that the surfaces would consist of a series of ridges running out from the center, with valleys in between. Thomas showed that the strength of the resulting field would increase toward the outside, compensating for the protons' relativistic increase in mass, and would also focus the protons so that they would stay in the vacuum chamber. Thomas's scheme was far too complicated for the techniques of the time, and it was ignored. Now we realize that he had anticipated the strong-focusing principle. Two Thomas-type cyclotrons are now under construction, one at Oak Ridge National Laboratory, the other at Berkeley. Both of them will produce protons and deuterons in the range of several hundred mev.

We can now understand an FFAG type of accelerator if we imagine that the radial scallops of the Thomas magnet are twisted into spiral ribs. (Is this the flamboyant style that presaged the end of the Gothic period?) The twisting introduces an additional kind of strong focusing. In fact, the idea grew out of strong focusing; only later was its similarity to the Thomas cyclotron recognized. The idea of FFAG has been exploited to the full by the workers of the MURA laboratory at Madison, Wis. They have imagined and computed (using two high-speed computing machines) all sorts of variations of the FFAG geometry, and have built several models that have successfully demonstrated the practicality of the scheme.

The advantage of the fixed-field design is twofold. First, it is easier to control a constant field than a varying one. Second, the fixed-field machines can be operated continuously, whereas the synchrotrons and synchro-cyclotrons must operate cyclically, or in pulses, a new cycle starting each time the field reaches its maximum value. Continuous operation means that more accelerated ions are produced per unit time; in other words, the beam has a higher intensity.

According to the MURA workers, the increased intensity that can be obtained with FFAG machines will make it possible to circumvent a serious limitation



USEFUL ENERGY in a collision depends on the motion of the particles after impact. Solid arrows at left represent energy of motion of bombarding particles. Solid arrows at right show energy of motion of the system after impact. Broken arrows indicate fraction of total energy available for desired reactions. Small dots are light particles; large dots, heavy ones. When like particles are made to collide head-on (*bottom*), all of their energy is available.

on accelerators which I have not mentioned as yet. This limitation concerns the amount of energy that is actually available to produce the reactions we are looking for.

When a high-energy ion from an accelerator strikes a stationary target particle, part of the energy goes into moving the target, and is wasted. It is as if we were trying to break a stone by hitting it with a hammer. To the extent that the hammer blow simply moves the stone, the energy is not available for breaking it. Now if the hammer is very light and the stone very heavy, we can see that the target will not move very far; almost all the energy of the hammer will go into breaking or chipping the stone. If we use a heavy sledge on a light pebble, most of the energy goes into moving the stone, and very little of it is available for breaking the stone. If the hammer and stone weigh the same, they will tend to move off together with half the speed of the incoming hammer; exactly half the energy will be available for breaking the stone.

It is the same with atom-smashing. But here relativity plays a particularly dirty trick, robbing us of most of the advantage to be gained by increasing the energy of the bombarding particles. We have seen that really high energies mean an increase in mass. Thus as we go up in energy we increase the weight of our "hammer" and lose a larger and larger fraction of its energy. At 1 bev a proton is already noticeably heavier than when it is at rest; when it hits a stationary proton, 57 per cent of the energy is wasted and only .43 bev is available for useful purposes. At 3 bev (the energy of the Brookhaven Cosmotron), the available portion is 1.15 bev; at 6 bev (the Berkeley Bevatron) the available portion is 2 bev; at 10 bev, 2.9 bev are available; at 50 bev, 7.5; at 100 bev, 10.5. We see that increasing the energy 100 times from one to 100 bev results in only a 20-fold actual gain.

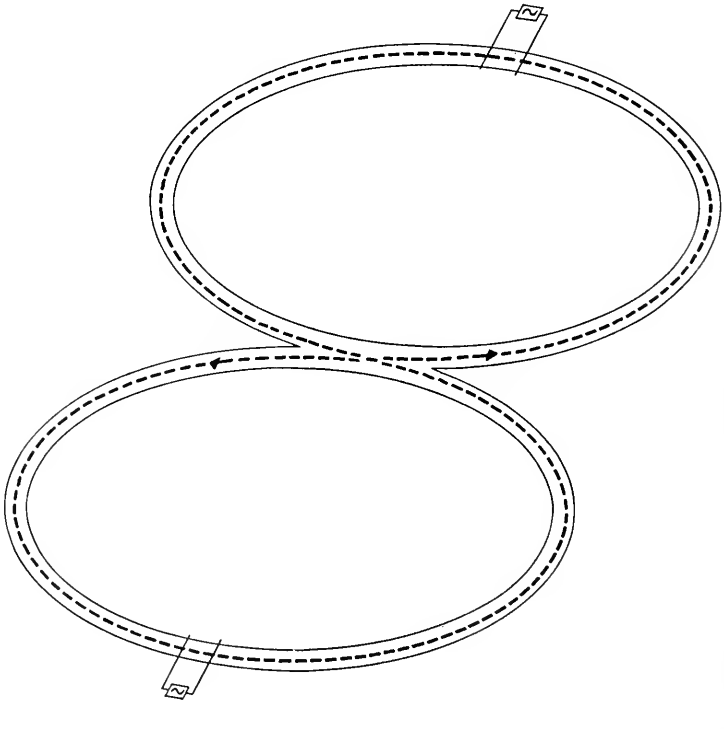
Suppose, however, that instead of firing a moving particle at a stationary one, we arrange a head-on collision between two high-energy particles. Then the mass increase is neutralized, and there is no tendency for the colliding particles to move one way or the other. All the energy of both of them is now available for the desired reactions. This is what the MURA designers propose.

They have envisaged a bold design, called "synchroclash," in which two 15-bev accelerators are placed so that their proton beams intersect and the particles collide with each other. This will yield an available energy of 30 bev, whereas in the case of a 30-bev proton colliding with a proton at rest only 6 bev would be available. In fact, to attain a useful energy of 30 bev in the ordinary way would mean using at least 500 bev. The success of the synchroclash idea turns on the intensity of the accelerator beams: there must be enough protons to make collisions reasonably frequent. The MURA proposal languished for several years, but interest in it seems to have revived. Perhaps the complicated orbits of the artificial satellites have had something to do with the new willingness to consider attempting the complicated orbits of FFAG.

Soviet Ideas

The Soviet designers have gone off in different directions. Veksler has been thinking of a scheme in which one approaches the ideal accelerator, namely one in which the accelerating field appears exactly in the vicinity of the ions but nowhere else. He envisages a small bunch of ions in a plasma (a gas of ions) exciting oscillations or waves in an electron beam. These waves are to act together coherently to give an enormous push to the ions being accelerated. If this is not clear to the reader, it is because it is not clear to me. The details have managed to escape most of us because of a linguistic ferrous curtain, but Veksler speaks of the theoretical possibility of attaining energies up to 1,000 bev. The proof of the idea must wait until it is put into practice. It should be remarked, however, that other wild schemes of Veksler, for example the synchrotron principle, are incorporated into most of our conventional accelerators today.

G. I. Budker of the U.S.S.R. has also presented some speculative ideas which have obviously been inspired by efforts to produce controlled thermonuclear reactions. Budker proposes an intense circular electron beam maintained by a weak magnetic guide field. The high current of the beam will cause it to "pinch" to a very small diameter because of its own magnetic field. The idea then is to use the very strong local magnetic field around the pinched beam as the guide field of a conventional accelerator [*see diagram on page 13*]. With an electron beam six meters in diameter one could expect to hold protons with an energy as high as 100 bev. Budker and his colleagues have constructed a special accelerator in which they have achieved a 10-ampere current of 3-mev



SYNCHROCLASH design would set two accelerators side by side so that their beams overlapped. Head-on collisions between particles would provide the maximum of useful energy.

electrons, and they expect to attain much higher currents and energies before long. It could well be that something really revolutionary will come out of this energetic work.

Our own thermonuclear program has inspired research on very strong magnetic fields [see "Strong Magnetic Fields," by Harold P. Furth et al.; SCIENTIFIC AMERICAN, February]. It seems likely that this development will find an application to the guidance of particles in multi-bev accelerators.

Electron Accelerators

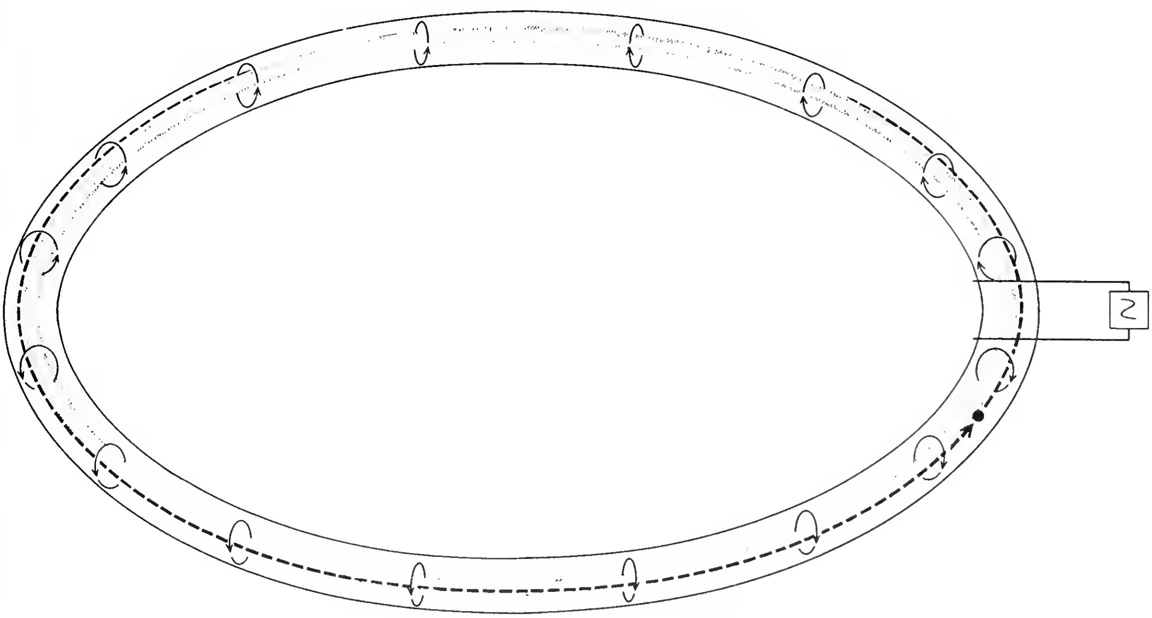
These new machines we have been discussing are proton accelerators, but there is vigorous activity in electron machines as well. We have already mentioned the Harvard-M.I.T. synchrotron which will attain 6 to 7.5 bev, and the half-dozen other smaller machines in the billion-volt range. The 220-foot linear electron accelerator at Stanford University has been on the scene for some time. Its energy has steadily increased so that it may now be used in experiments at 600 mev. We expect to welcome it to the 1-bev club before long.

The linear machine is significant because there is a special difficulty in reaching high energy with electron synchrotrons. When electrons are made to travel on a curved path at high speeds they give off strong electromagnetic radiation. The effect is easily visible to the naked eye; the luminous horizontal beam on the cover of this issue of SCIENTIFIC AMERICAN is synchrotron radiation. The difficulty is that this radiation can represent a substantial loss of energy, and it increases rapidly as the energy of the machine goes up. In the Harvard-M.I.T. synchrotron the amount of energy radiated is almost prohibitive (about 10 mev per turn at 7.5 bev). To reach higher energies, say 20 bev, the Stanford group has been thinking in terms of a linear accelerator, which does not have this radiation difficulty because its particles do not move in a circle. Such a machine might be as much as three miles long.

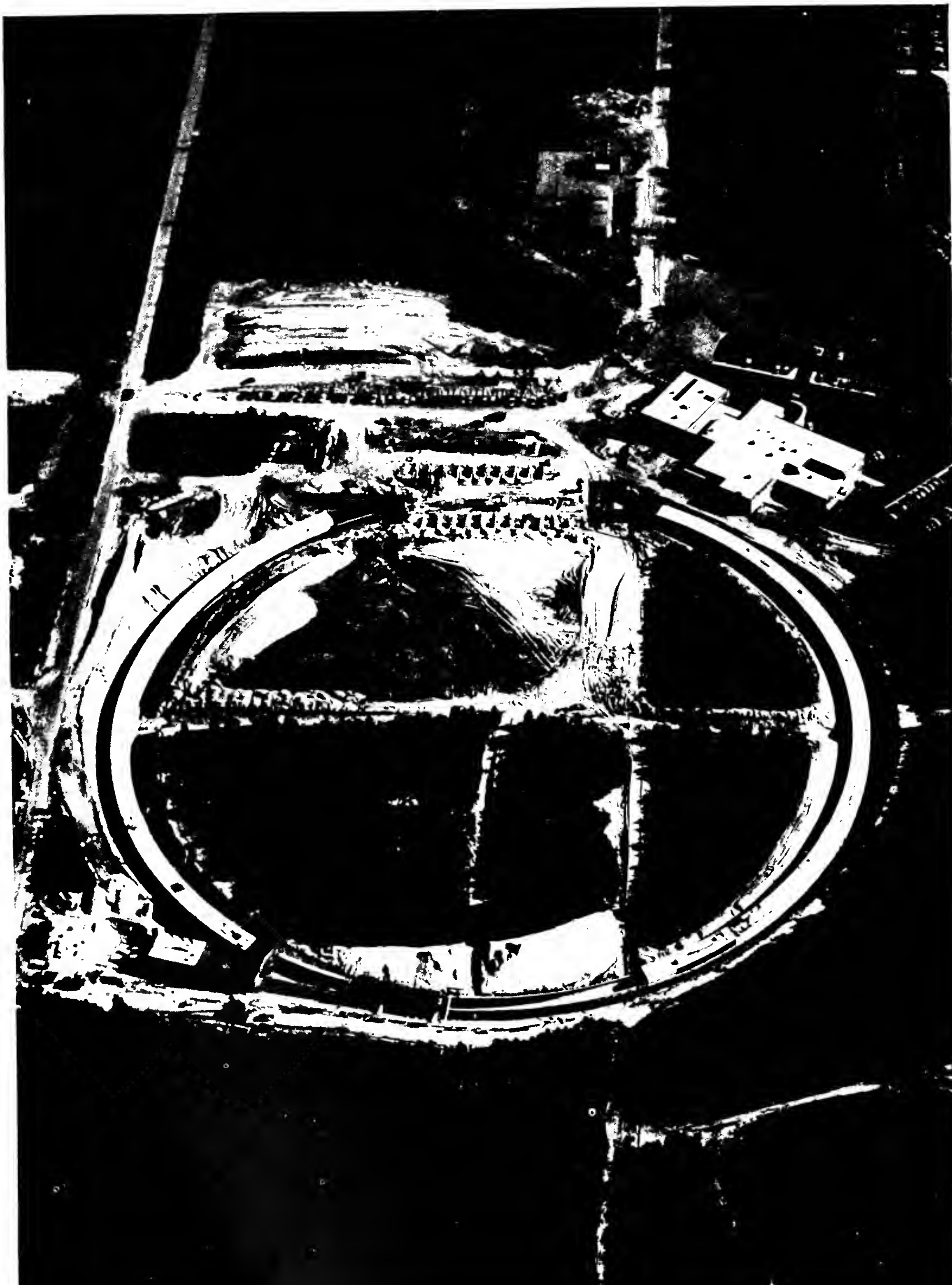
I am not convinced that the limit of electron synchrotrons has been reached. Indeed, it is not difficult to imagine a 50-bev electron synchrotron. The radiation problem would be solved by reducing the curvature of the electron beam, that is, by increasing its radius to, say, half a mile. I believe that the upper limit of the electron synchrotron may be as high as 100 mev.

While we are "thinking big" we should not forget Enrico Fermi's proposal to ring the earth with a vacuum tube and, using the earth's magnetic field, obtain 100,000 bev. For that matter, now that artificial satellites are commonplace, we might put up a ring of satellites—each containing focusing magnets, accelerators, injectors and so on—around the earth. Something like a million bev could be expected from this accelerator, which we might as well call the lunatron. At the very least such a device will eliminate the need for vacuum pumps, since it will be outside the atmosphere.

Villard de Honnecourt and later Viollet-le-Duc have left us detailed accounts of the builders of cathedrals and of their methods. It seems to be pretty much the same story then and now. The designer of the cathedral was not exactly an architect, nor is the designer of an accelerator exactly a physicist. Both jobs require a fusion of science, technology and art. The designers of cathedrals were well acquainted with each other; the homogeneity of their work in different countries is evidence of a considerable interchange of information. The homogeneity of accelerator design demonstrates the same interchange today. Our medieval predecessors were only human; one gets the definite impression that they were subject to petty jealousies, that occasionally there was thievery of ideas, that sometimes their motivation was simply to impress their colleagues or to humiliate their competitors. All these human traits are occasionally displayed by their modern counterparts. But one also gets a strong impression of the excitement of those mighty medieval creators as they exulted in their achievements. This sense of excitement is no less intense among modern nuclear physicists.



PINCH EFFECT might be used to provide a magnetic guiding field for an accelerator, thus eliminating the heavy magnet. The dotted ring is a pinched plasma. Its magnetic field, which is shown by colored lines, would act to hold particles near its outer edge.

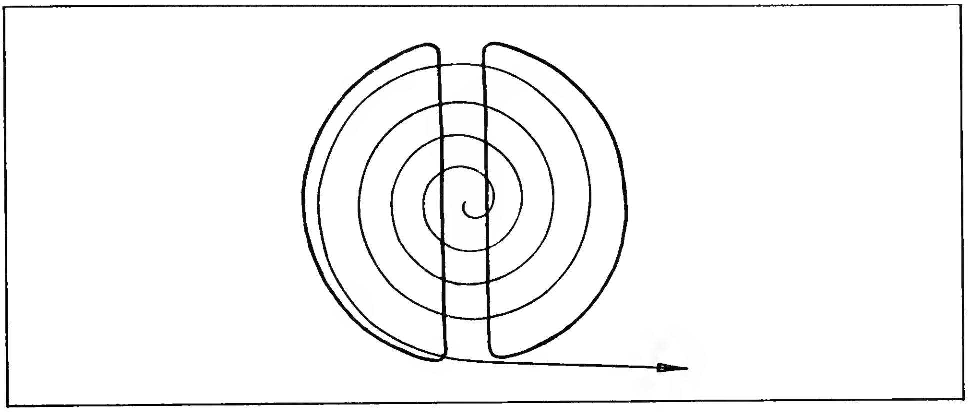


HUGE PROTON SYNCHROTRON under construction at Brookhaven National Laboratory is photographed from the air. Circular tunnel housing its doughnut is 840 feet in diameter. This machine will produce particles of 25 to 30 billion electron volts (bev).

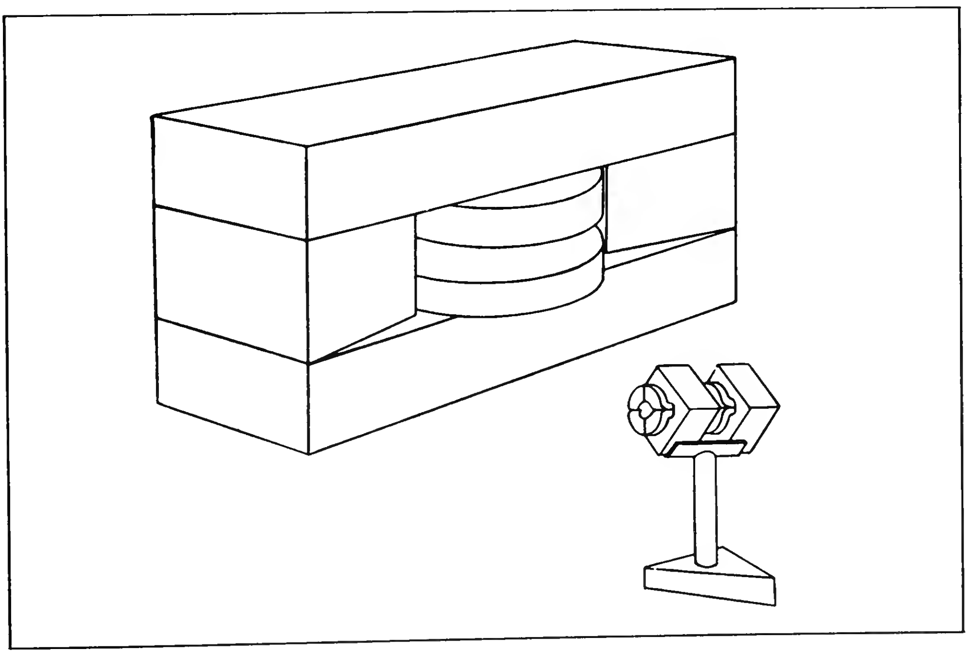
9 The Cyclotron As Seen By...

David L. Judd and Ronald G. MacKenzie of the Lawrence Radiation Laboratory, University of California, Berkeley

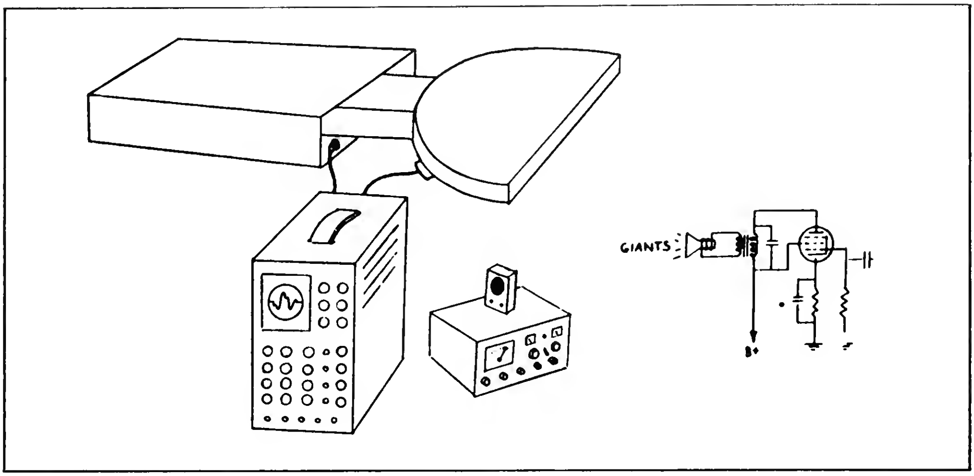
The cartoons were prepared to accompany Dr. Judd's keynote address at the International Conference on Isochronous Cyclotrons at Gatlinburg, Tennessee, May 1966.



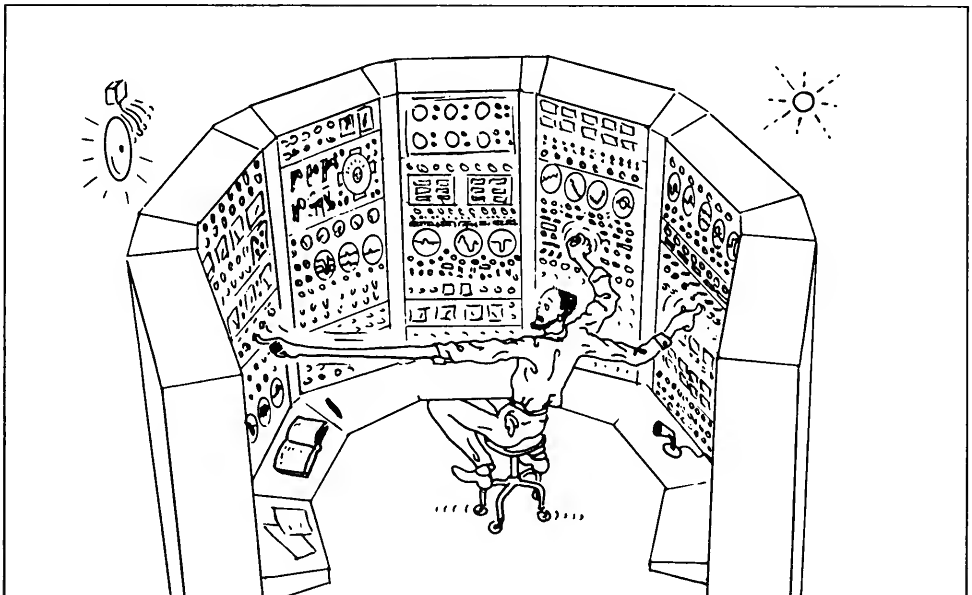
The Cyclotron as seen by the inventor



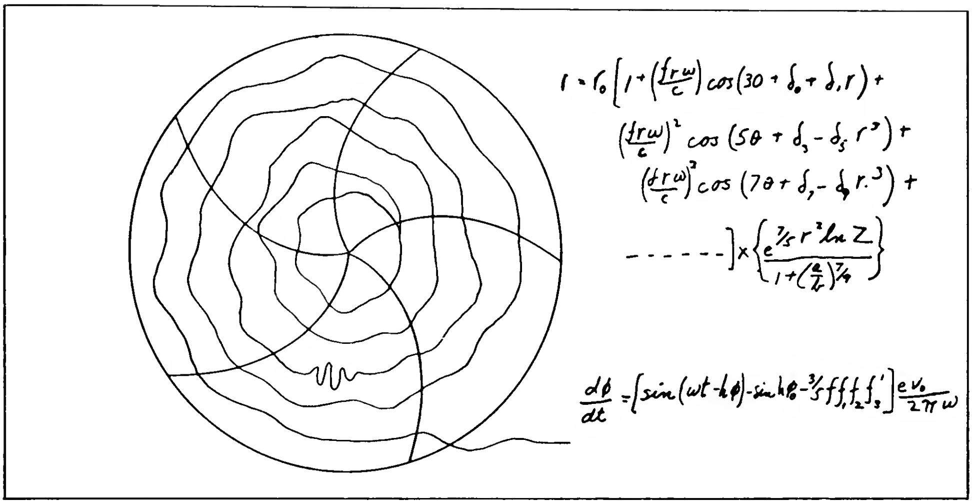
The Cyclotron as seen by the Mechanical Engineer



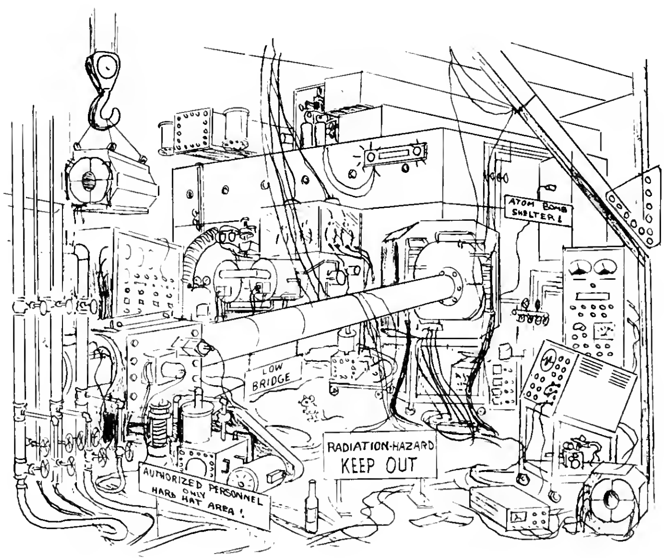
The Cyclotron as seen by the Electrical Engineer



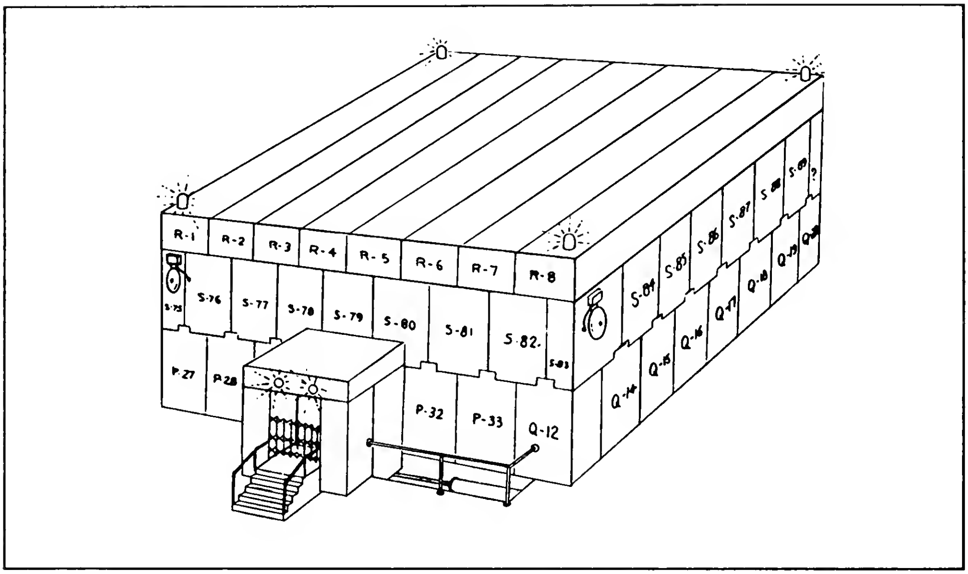
The Cyclotron as seen by the operator



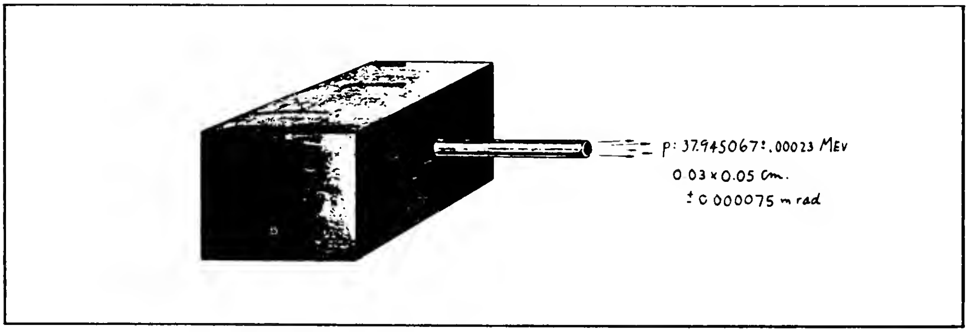
The Cyclotron as seen by the Theoretical Physicist



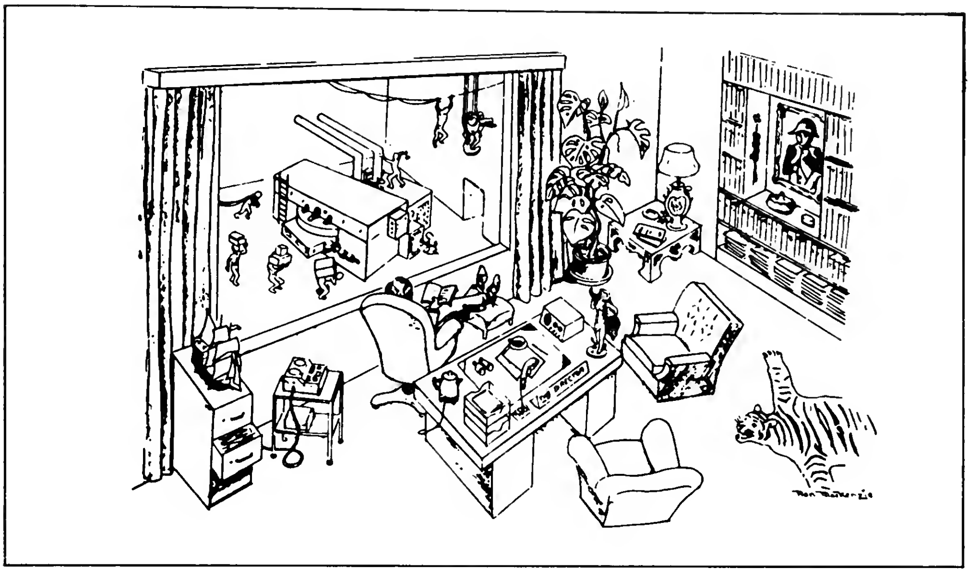
The Cyclotron as seen by the Visitor



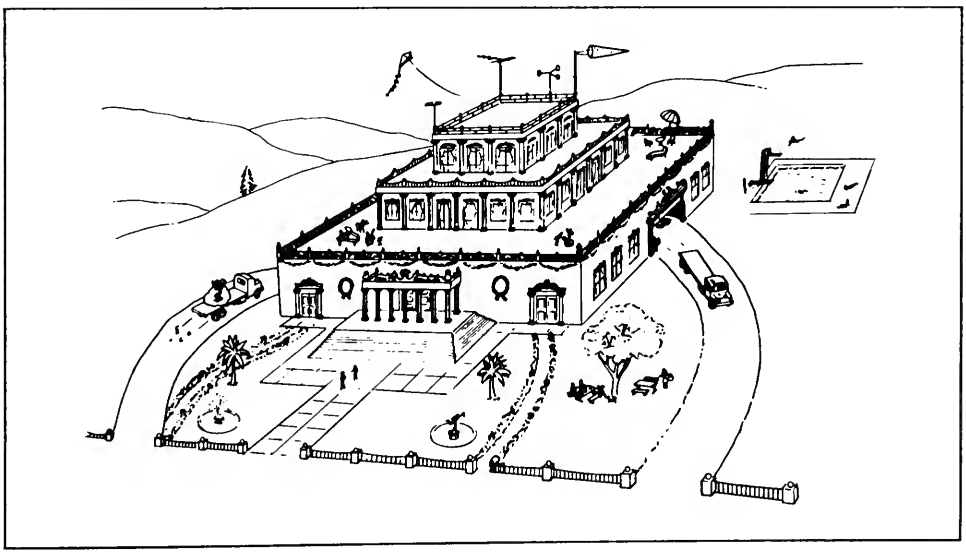
The Cyclotron as seen by the Health Physicist



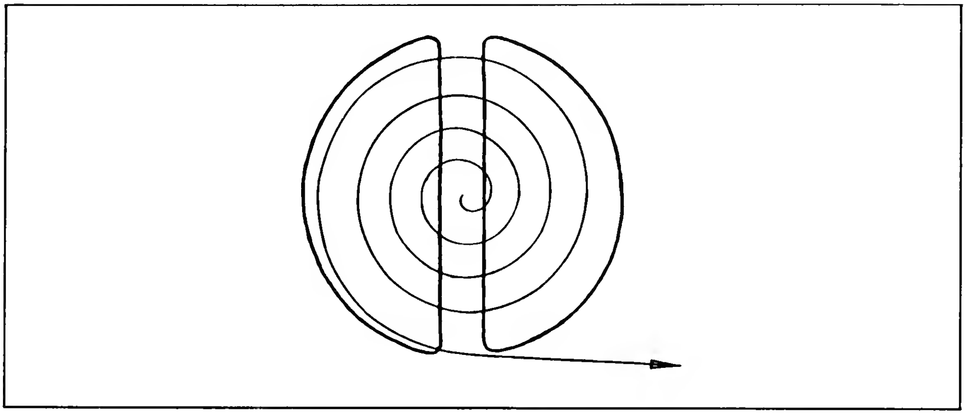
The Cyclotron as seen by the Experimental Physicist



The Cyclotron as seen by the Laboratory Director



The Cyclotron as seen by the Government Funding Agency



The Cyclotron as seen by the student

CERN (Conseil Européen pour la Recherche Nucléaire) is an installation created to pool the finances and talents of many European nations. Physicists come there from all over the world to work together in high-energy physics research.

10 CERN

Jeremy Bernstein

Article published originally in *The New Yorker* in 1964.

SHORTLY after the Second World War, when the normal international life of science was resumed, a physicist who had just listened to several hours of technical lectures at a large conference remarked that the international language of physics had become a combination of mathematics and broken English. Today, almost all scientific journals, including the Russian—and even the Chinese journals, such as the *Acta Mathematica Sinica*, and *Scientia Sinica*, published in Peking—give at least the title of each article, and often an abstract, in English. From the title and the equations and the graphs, a specialist in the field can usually reconstruct the general theme of the article. The exchange of articles and journals among scientists of different countries is one of the oldest and best traditions of science. It goes on independently of the political climate. During the darkest days of the Stalinist period in Russia, scientific papers went back and forth across the Iron Curtain, and Western physicists could follow the work of such Russians as Lev Landau (the most distinguished Russian theoretical physicist, who won the Nobel Prize in 1962), despite the fact that he was under house arrest in Moscow, in part because of his liberal ideas and in part because he is a Jew.

With the death of Stalin and the relaxation of some of the tensions between East and West, it became possible for scientists to travel in and out of the Eastern countries. The so-called Rochester Conference in High-Energy Physics (it gets its name from the fact that the first seven conferences, starting in 1950, were held in Rochester, New York) now meets one year in the United States, one year in Geneva, and one year—indeed, last summer—in the Soviet Union. Several American universities have regular exchange programs with Soviet universities, and it is no longer a novelty to find a Russian physicist giving a series of lectures in an American university, and vice versa.

The ultimate in international scientific coöperation is, of course, the international scientific laboratory, in which scientists of many countries can actually work together. In fact, it is becoming increasingly clear that such laboratories are not only desirable but necessary. Research in a field like high-energy physics—in a way, the most basic of all the sciences, since it is the study of elementary particles, the ultimate constituents of all matter—has become so expensive that many people have come to believe that pursuing it as a purely national enterprise is difficult to justify. A recent editorial in the New York *Times* pointed out that "high-energy physicists . . . use the most elaborate and most expensive equipment employed in any branch of terrestrial basic research," and went on to say, "These are the particle accelerators, which today cost tens of millions of dollars each, and which will in the future be priced in the hundreds of millions. The Atomic Energy Commission's operating and construction costs in this field are already expected to aggregate $165 million in the next fiscal year, and one authoritative estimate places the annual bill by the end of the next decade at $370 million, reaching $600 million by 1980. . . . Nuclear physicists are already talking about far more powerful—and much more expensive—atomic-research instruments. The case for building these machines is an impressive one, but the case for building them only with the resources of one country is not convincing."

The editorial concluded by pointing out that there already exists an excellent working example of an international atomic laboratory; namely, CERN (standing for Conseil Européen pour la Recherche Nucléaire), which is operated jointly by almost all the Western European countries and is situated in the Swiss town of Meyrin, a suburb of

Geneva that is almost on the French frontier. CERN itself sprawls along the frontier, and recently, when it needed room for expansion, the French government gave it a ninety-nine-year lease on a hundred acres of French land, matching the hundred acres of Swiss territory that the center now occupies. This makes CERN the only international organization that actually straddles a frontier. Its facilities include two accelerators (the larger, a proton synchrotron, accelerates protons to energies up to twenty-eight billion electron volts, and shares with its slightly more powerful twin, the alternating-gradient synchrotron at the Brookhaven National Laboratory, on Long Island, the distinction of being the largest accelerator now operating), several electronic computers, and a vast collection of bubble chambers, spark chambers, and other paraphernalia necessary for experimenting with the particles produced in the accelerators—to say nothing of machine shops, a cafeteria, a bank, a travel agency, a post office, a large library, and a multitude of secretarial and administrative offices. It costs about twenty-five million dollars a year to run. This money is contributed by thirteen European member states—Austria, Belgium, Great Britain, Denmark, France, Greece, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland, and West Germany. Neither the United States nor Russia is eligible to become a member, since neither is "Européen," but there are Americans and Russians who work at CERN. An exchange agreement exists between CERN and DUBNA, a similar laboratory near Moscow, where physicists from the Iron Curtain countries and China work together. Each year, DUBNA sends two or three physicists to CERN for several months at a time. American physicists at CERN have been supported by sabbatical salaries, by fellowships like the Guggenheim and the National Science Foundation, or by money from Ford Foundation grants (totalling a bit over a million dollars) that were given to the laboratory explicitly for the support of scientists from non-member countries. (The grants have now been discontinued, following the Ford policy of "pump-priming," and the laboratory is looking for other sources of money.) There are usually twenty or twenty-five Americans at CERN. In addition, the laboratory has contingents of Japanese, Indians, Poles (a very active and scientifically strong group of about a dozen), Yugoslavs, Turks, Israelis (there is an exchange agreement with the Weizmann Institute, in Rehovoth), and Hungarians. All the permanent personnel at CERN—about sixteen hundred people, of whom about three hundred are physicists and engineers—are drawn from the member states. (Their average age is thirty-two.) As one might imagine, all this produces a tutti-frutti of languages, national types, political attitudes, and social mannerisms, and everyone accepts and enjoys the chaos of national flavors as part of the working atmosphere of the laboratory. As an American physicist and a perennial summer visitor to CERN, I have had fairly typical experiences there. This past summer, I worked with an Italian physicist in an attempt to extend some work done by a German-born American physicist who was visiting the laboratory on a Guggenheim Fellowship. This work was itself an extension of another Italian physicist's work, which, in turn, was based on the work of an American physicist who is a frequent visitor to CERN. (I also helped a Yugoslav physicist with the English translation of a short book written by a well-known Russian physicist whom I met when he visited CERN to attend the Rochester Conference of 1962, which was held in Geneva.) My working language with the Italian physicist was English (and, of course, mathematics). Most of the people at the laboratory are polylingual. All scientific lectures are given in English, and almost all the technical personnel have a good command of the language. However, the language one hears most often is French; the secretaries, postmen, bank clerks, mechanics, and telephone operators speak it among themselves, and so do many of the European physicists. Secretaries must be able to type technical manuscripts in English, since almost all the publications that come out of CERN

each year (several hundred of them) are in that language.

Because nuclear physics has become so closely associated (at least in the public mind) with its military applications, many people have wondered how a laboratory that intermingles physicists from the East and the West—and, indeed, from all over the world—can possibly operate without running into all sorts of problems of military security and national secrecy. The answer is that nuclear physics is a very broad subject. It ranges from the study of nuclear energy—fission, fusion, reactors, and the like—to the study of the interior structure of the nucleus, and even to the study of the structure of the very neutrons and protons and other particles that compose the nucleus. This latter study is the frontier of modern physics. Because high-energy particles are necessary in order to probe deeply into the interior of the nucleus, this branch of physics is called "high-energy," as opposed to "low-energy," or "classical"—"classical" in that the laws governing the behavior of the nuclei in, say, the fission process in a reactor are now pretty well understood, and have been for some time. The military and technological applications of nuclear physics are based on these latter laws, whereas the study of the interior structure of the nucleus has no technological applications at present; more than that, it is difficult now to imagine any such applications in the future. However, the example of Einstein's special theory of relativity—one of the most abstract theories in physics—which has been the basis of the entire development of nuclear energy, shows that theoretical speculations that may at the moment seem far removed from reality can very quickly change all of technology.

THE very fact that high-energy physics does not have military applications was among the reasons it was chosen as the discipline for an international laboratory. In the late nineteen-forties, when a number of prominent physicists—including the late H. A. Kramers, of Holland; Pierre Auger and Francis Perrin, of France; Edouardo Amaldi, of Italy; and J. Robert Oppenheimer, of the United States—began informally discussing the prospects for creating an international laboratory in Europe, they set out to look for a field that would be sufficiently close to recent developments in atomic energy for European governments to be interested in supporting the project financially, and yet far enough removed from immediate applications of atomic energy for military security not to be a problem. They also realized that it would be necessary to engage the support of the European diplomats who were then promoting attempts to create a United Europe. One of the most influential of these diplomats was François de Rose, of France. (He is now the French Ambassador to Portugal.) De Rose became interested in the possibilities of atomic research just after the war, and in 1946 he met with Oppenheimer in New York at the United Nations Atomic Energy Commission. Out of the resulting friendship between the two men an important link developed between the scientific and diplomatic communities. Dr. L. Kowarski, a French nuclear scientist and one of the pioneers of CERN, has written a semi-official history of the origins of the laboratory, in which he notes:

> The first public manifestation of this new link occurred in December, 1949, at the European Cultural Conference held in Lausanne. A message from Louis de Broglie [de Broglie, the most distinguished French theoretical physicist of modern times, was awarded the Nobel Prize in 1929 for his work on the wave nature of electrons] was read by Dautry [Raoul Dautry was at that time the administrator of the French Atomic Energy Commission and one of the leaders of the movement for a United Europe], in which the proposal was made to create in Europe an international research institution, to be equipped on a financial scale transcending the individual possibilities of the member nations. . . . At that time [a dilemma] was besetting the scientists' aspirations: atomic energy was attracting public readiness to spend money, but atomic energy invited security-mindedness and separatism. The way out of the dilemma was clear enough. The domain of common action should be chosen so as not to infringe directly the taboos on uranium fission, but [to be] close enough to it so as to allow any successes gained internationally in the per-

mitted field to exert a beneficial influence on the national pursuits.

The ultimate choice—high-energy physics—was a perfect compromise; although it is a branch of nuclear physics, it is one that is far removed from military applications.

In June of 1950, the American physicist I. I. Rabi initiated the first practical step toward the creation of such a pan-European laboratory. As a member of the United States delegation to UNESCO he was attending the UNESCO conference held that year in Florence. Speaking officially on behalf of the United States, he moved that UNESCO use its good offices to set up a physics laboratory (he had high-energy physics in mind) with facilities that would be beyond those that any single European country could provide, and that would be comparable to the major American facilities at Brookhaven and Berkeley. It was an important step, because it placed the prestige and influence of American science behind the project. The implementation of Rabi's motion became the work of Pierre Auger, of France, a distinguished physicist who was the UNESCO scientific director. As a result of his efforts, various cultural commissions of the French, Italian, and Belgian governments donated about ten thousand dollars for a study program, and CERN was under way. (In the course of the discussions held at that time, Rabi stressed the desirability of not having any nuclear reactors at CERN, since they have both military and commercial applications—and, in fact, there are none.) Dr. Kowarski writes:

Two objectives were suggested: a longer-range, very ambitious project of an accelerator second to none in the world [this resulted in the construction of the proton synchrotron, which was completed in 1959] and, in addition, the speedy construction of a less powerful and more classical machine in order to start European experimentation in high-energy physics at an early date and so cement the European unity directed to a more difficult principal undertaking.

At the end of 1951, an organizational meeting was held in Paris; all the European members of UNESCO were invited, but there was no response from the countries of Eastern Europe. Then, at a meeting held in Geneva early in 1952, eleven countries signed an agreement pledging funds and establishing a provisional organization. There was something of a tug-of-war among the member countries to decide where the new laboratory should be built. The Danes, the Dutch, the French, and the Swiss all had suitable territory for it, but in the end Geneva was chosen, partly because of its central location, partly because of its long tradition of housing international organizations (there are, for example, all sorts of multilingual elementary schools in the city)—and, it is said, partly because some of the physicists involved in the decision were avid skiers. The Swiss government gave, free, the site near Meyrin, and in June, 1953, the Canton of Geneva formally ratified, by popular referendum, the government's invitation to CERN to settle there; in addition, the laboratory was given the same political status as that of any of the other international organizations in Geneva. At the same time, a formal CERN Convention was prepared for the signature of the member states, which then numbered twelve; Austria and Spain joined later, and Yugoslavia, an original signatory, withdrew in 1962, because of a lack of foreign currency. Article II of the Convention stipulates: "The Organization shall provide for collaboration among European States in nuclear research of a pure scientific and fundamental character, and in research essentially related thereto. The Organization shall have no concern with work for military requirements, and the results of its experimental and theoretical work shall be published or otherwise made generally available." The Convention also set up a formula for CERN's financial support. Roughly speaking, each member nation pays each year a certain percentage (a fraction of one per cent) of its gross national product. This means, in practice, that Great Britain, France, and West Germany pay the largest shares. The CERN Council, the governing body of the laboratory, was set up, with two delegates

from each country—one a scientist and the other a diplomat, like de Rose. The Council meets twice a year to pass on such matters as the budget and the future development of the laboratory. (During my last visit to CERN, there was a Council meeting in which the question of constructing a still larger international machine—a machine capable of accelerating protons to three hundred billion electron volts, or about ten times the capacity of the present machine—was discussed.) The Council also, by a two-thirds majority, appoints the Director-General of the laboratory. The Director-Generalship of CERN is a very complex job, and few people are really qualified for it. In the first place, the Director-General can have no special national bias. As the Convention puts it, "The responsibilities of the Director and the staff in regard to the Organization shall be exclusively international in character." In the second place, the Director must clearly be a physicist, for, among other things, he must decide which of various extremely expensive experiments the laboratory should concentrate on. The first Director, chosen in 1954, was Professor Felix Bloch, of Stanford University—a Swiss by origin and a Nobel Prize winner in physics. Professor Bloch returned to Stanford in 1955 and was succeeded by C. J. Bakker, a Dutch cyclotron builder. (Professor Bakker was responsible for the construction of the cyclotron, the smaller of the accelerators at CERN.) He held the post from 1955 to 1960, when he was killed in an airplane accident on his way to Washington, where he had intended to deliver a report on the operation of the large accelerator, the proton synchrotron, which had gone into operation in 1959.

IF any one individual was responsible for the successful construction of the large accelerator, it was John B. Adams, an Englishman, who took over the Director-Generalship on Bakker's death. Adams was born in 1920 in Kingston, Surrey, and received his education in English grammar schools. At eighteen, he went to work for the Telecommunications Research Establishment, and when the war broke out he joined the Ministry of Aircraft Production. He had received some training in electronics with the Telecommunications Establishment, and in the M.A.P. he became involved with the problem of installing the first radar in fighter planes. It soon became evident that he had a gift both for engineering and for the complex job of directing a large technical project. In fact, the war produced a whole generation of young scientists and engineers who not only were technically competent but had acquired considerable practical experience in running large-scale and costly scientific enterprises. These men moved readily into the various atomic-energy programs that were started after the war, and Adams joined the nuclear laboratory at Harwell, the principal British center for experimental work in nuclear physics. At this time, the people at Harwell were beginning work on a hundred-and-seventy-five-million-volt proton accelerator, and Adams became an important member of the project. The machine was finished in 1949, and Adams spent the next three years working on the design of special radio tubes needed in connection with accelerators. Then he was released by the Ministry of Supply to go to Geneva and join the new accelerator project at CERN.

By that time, the CERN group, which had been at work since 1951, had inherited a technological windfall in the way of accelerator design. A particle accelerator can accelerate only those particles that carry an electric charge. Advantage is taken of the fact that when a charged particle passes through an electric field it is accelerated by the force that the field exerts on it. In modern accelerators, transmitting tubes generate the electromagnetic fields, in the same way that radio transmitters generate radio waves. These accelerating stations are placed at intervals along the path of the particles in the machine, the simplest arrangement being along a straight line. This layout results in what is called a linear accelerator, or LINAC. The particles move faster and faster in a straight line and

are finally shot out the other end into a target of some sort. The energy that such particles can acquire is limited by the length of the straight line, as well as by the power of the transmitters. At Stanford University, there is a nearly completed straight-line accelerator, known among physicists as "the monster," that will accelerate electrons over a path almost two miles long; the emerging electrons will have an energy of about twenty billion electron volts. Most accelerators, however, are circular. The accelerating stations are arranged along the perimeter, and as the particles go around and around they acquire more energy in each orbit. This arrangement saves space and greatly reduces the number and size of the accelerating stations. The problem that naturally arises is how to maintain the particles in circular paths while they are being accelerated, since a particle will move in a circle only if a force acts on it to keep it from flying off at a tangent. In circular accelerators, this force is supplied by electromagnets. The magnets are deployed along the path of the particles, and the magnetic fields they produce hold the particles in orbit. The drawback to this system is that the more energy a particle acquires, the more strongly it resists staying in a circular orbit and the larger the magnet required to keep it so. In fact, as the postwar accelerators became more and more powerful, the size of their magnets began to get out of hand. The Brookhaven cosmotron, a proton accelerator producing protons with an energy of three billion electron volts, has a magnet of four thousand tons; the Berkeley bevatron, with six-billion-electron-volt protons, has a magnet weighing ten thousand tons; and, most striking of all, the Russian phasotron at DUBNA, which produces protons of ten billion electron volts, has a magnet weighing thirty-six thousand tons.

This was where things stood in 1952, when the CERN group planned to make an accelerator of at least ten billion electron volts. By using a somewhat modified and more economical design than the one for the DUBNA machine, the new accelerator could have been made with a magnet weighing from ten to fifteen thousand tons, but even this seemed monstrous at the time. That year, however, a group at Brookhaven consisting of E. Courant, M. S. Livingston, and H. Snyder, in the course of solving a problem put to them by a group of visiting accelerator experts from CERN, invented a principle of magnetic focussing that altered the situation completely. (It turned out later that their method had been independently invented a few years earlier by an American-born Greek named N. Christofilos, who was employed in Greece selling elevators for an American firm and was a physicist in his spare time. Christofilos had sent a manuscript describing his invention to Berkeley, where it was forgotten until news of the work at Brookhaven reminded somebody of it. Christofilos is now at the Livermore Laboratory of the University of California.) The magnet in a circular accelerator not only bends the particle trajectories into circles but applies a force that focusses the beam and keeps it from spreading out indefinitely as it goes around and around. The magnets in the old machines could supply only very weak focussing; thus the beam was pretty thick, and the vacuum pipe it circulated in and the magnet surrounding it also had to be large. (At Berkeley, a man can crawl through the vacuum chamber.) It was known, however, that magnets could be made that would give much stronger focussing forces, but only in one direction at a time; that is, if the beam were kept confined horizontally it would expand vertically, and vice versa. What the Brookhaven people found was that if an accelerator magnet ring was built up of alternate sections that provided strong focussing and defocussing forces, the net result was a focussing much stronger than anything that had previously been achieved. (In the new machines, the beam can be contained in a vacuum pipe only a few inches in diameter.) This meant that the magnets could be much smaller in size, with a great saving of weight, power, and cost. The CERN magnetic system weighs only three thousand tons, although the proton energies achieved are

nearly three times those generated by the old Russian machine, which had a magnet weighing over ten times as much. The focussing works so well that the final beam of particles, which consists of about a thousand billion protons per second, is only a few millimetres wide when it emerges from the machine. The ring around which the protons race is about two hundred metres in diameter. The protons are injected into the main circular track by a small linear accelerator, and in the single second that they remain in the machine they make about half a million revolutions. The entire ring must be kept at a fairly high vacuum, since otherwise the protons would knock about in the air and be scattered. There is also a delicate question of timing. The accelerating fields must deliver a kick to each bunch of protons at just the right instant in its orbit. As the protons move faster and faster, approaching the speed of light, the synchronization of the fields and the particles must be constantly changed. However, according to Einstein's special theory of relativity, no particle can go faster than light, so that near the end of the cycle the protons will be gaining energy but not speed (the particles, again according to the relativity theory, get heavier and heavier as they move faster and faster), which simplifies the timing problem somewhat. Indeed, high-energy-accelerator design, which uses the theory of relativity extensively, and which clearly works, is one of the best-known tests of the theory itself. That all these factors, complex as they are, can be put together to make a reliably operating machine is an enormous triumph of engineering physics.

Needless to say, the Brookhaven people were eager to build a machine operating on the principle they had invented. However, the cosmotron had only recently been finished, and they could not get immediate support for the construction of an even larger machine—especially one that would use a principle still untested. The CERN people, however, were in a much more advantageous position, and in 1953 they began designing the laboratory's present machine, the CPS (CERN proton synchrotron). About six months later, influenced partly by the progress at CERN, the Brookhaven people got under way with the construction of a similar but slightly larger machine—the AGS, or alternating-gradient synchrotron. A friendly race developed between the two groups, with CERN finishing in November, 1959, and Brookhaven about six months later.

In order to construct the CERN accelerator, Adams gathered around him a superb international team of engineers and physicists interested in accelerator construction. Not only is he a brilliant engineer himself but he has the ability to organize other engineers into effective groups with physicists, so that very new ideas can be effectively realized on an industrial scale. In fact, working on the accelerator at CERN came to be a considerable distinction for an engineer, and CERN got almost the pick of the European engineers, even though the laboratory could not compete financially with the salaries that were being offered by European industry. The machine was so well designed that it worked better than had been generally anticipated. It became available to the physicists at CERN early in 1960, and Adams stepped into the gap caused by Bakker's death to become Director-General of the laboratory for a year. He also received an honorary degree from the University of Geneva, which he accepted on behalf of the group that had worked with him. He is now back in England directing a laboratory that is studying the problem of controlling nuclear-fusion energy for general application. (Nuclear-fusion energy arises when nuclear particles are fused to make a heavier nucleus. The heavier nucleus actually weighs less than the sum of its parts, and—again according to Einstein's relativity theory—the excess weight is liberated as energy. The hydrogen bomb is an unfortunate application of this principle.)

THE present Director-General of CERN is Professor Victor F. Weisskopf, who was given leave of absence from M.I.T. to take over from Adams in 1961. Professor Weisskopf,

whom I got to know when I was a student at Harvard in the nineteen-fifties, was born in Vienna, so although he is an American citizen, he can be counted as a European. He is one of the world's leading theoretical physicists, as well as one of its most likable. A large, friendly man, he is known to almost everybody at CERN as Viki, and despite a recent and very serious automobile accident he remains a devoted skier and hiker. This past summer, I had several talks with him about the development of CERN. One of the most interesting observations he made had to do with the evolution of the present generation of European physicists. At the end of the war, he said, European physics, which had been the finest in the world, was greatly damaged. Many of the best European physicists were more or less permanently settled in either England or the United States and had no desire to come back to Europe and relive a very unpleasant experience. In particular, the tradition of experimental physics, which requires complicated equipment, had greatly suffered on the Continent during the years of deprivation. Consequently, when the big accelerator at CERN was ready, there was a shortage of highly trained European experimenters to use it. On the other hand, the war had greatly strengthened physics in the United States, not only because so many Europeans had come here to live but because physicists had been working all through the war at places like Los Alamos on subjects that were not entirely dissimilar to their peacetime research. Thus, the postwar generation of American physicists was highly trained and ready to continue along the line of research that had made the development of high-energy physics the frontier of physics. (Many of the early research papers written at CERN during this period were done by Europeans in collaboration with Americans at the laboratory, some of whom had been born in Europe and were back on visits.) Of even greater importance, most of the European physicists who currently have important positions at CERN spent time in the United States, where they received training in the then novel techniques of experimental physics. As Weisskopf pointed out, a new generation of excellent and inventive physicists has by now grown up in Europe. They are producing scientific work at the forefront of modern physics that is of the first quality and the equal of anything being done in the United States or Russia. These physicists are now training young Europeans, to say nothing of American postdoctoral visitors. Originally, some European university professors were opposed to the creation of CERN on the ground that it would draw too many scientists away from the universities at a time when there was a desperate shortage of them. Weisskopf remarked that it has worked out almost the other way—that European physicists have come to Geneva for a few years of advanced training and then gone back to their own countries to teach and do research in universities. In fact, according to many of the young European physicists I have spoken to, it is now quite hard to find good jobs in European universities, and CERN offers an opportunity to continue working until a suitable position opens up somewhere.

FOR me, one of the most interesting experiences at CERN was the contact with some of the Russian physicists at the laboratory. As a rule, the Russians who come to Geneva are about equally divided between experimental and theoretical physicists. Because high-energy experimental physics is done by teams, the experimental physicists join a group of other experimenters, while the theorists work pretty much alone. As it happened, one of the Russian experimenters—Vitaly Kaftanov, from the Institute for Experimental and Theoretical Physics, in Moscow—was working on an experiment that was of special interest to me, since I had been studying some of its theoretical implications. This experiment—one of the most elaborate and active at CERN—involves the study of reactions induced by neutrinos. The neutrino is a marvellous particle. It is almost impossible to detect directly, for it has no charge and no mass, and it interacts very weakly with ordinary matter. Indeed, someone has estimated that if one took a single neutrino produced in the accelerator at CERN or the

one at Brookhaven (where the first high-energy neutrino experiments were done) and shot it through a layer of lead about as thick as the distance from here to Pluto, it would undergo only one collision during its entire passage. Fortunately, however, the experimenter is not limited to one neutrino; an accelerator produces millions of them a second, and some are bound to make a collision in a target of reasonable size. These collisions produce particles that *can* be seen, so that neutrino reactions can be studied. Since the collisions are so rare, the whole experimental area must be carefully shielded from cosmic rays and other annoying background that could be confused with the few events that one is looking for. In the experiments both at CERN and at Brookhaven, this required literally thousands of tons of heavy shielding material. (The shielding in the Brookhaven experiment was made from the remnants of an obsolete battleship, while at CERN it consists of steel ingots lent to the laboratory by the Swiss government from its strategic stockpile.) At both CERN and Brookhaven, neutrino events have been successfully detected; in fact, in the Brookhaven experiment it was first shown that there are two quite distinct species of neutrino. Until that experiment, the neutrino was generally taken to be a single, unique particle (although there were some theoretical conjectures to the contrary). The fact that precision experiments can now be done with neutrinos is a very important breakthrough in the technology of experimental physics, and it is only natural that a physicist like Kaftanov is eager to work on the project.

Kaftanov, who is married and has a young son, first came to CERN alone. This past summer, he was joined by his family. He has a warm, friendly personality and a good command of English. (He told me that when he was young his parents agreed to allow him to give up music lessons, which he hated, on condition that he study English.) Many of our conversations concerned the progress of the experiment, but as we got to know each other better we talked a good deal about a physicist's life in the United States and in Russia. In his country, physicists and engineers are at the very top of the social and economic scale, and the disciplines themselves are characterized by a highly didactic style. There is a great deal of sharp, sometimes quite personal criticism at all levels. Among European physicists, by contrast, there is still some feeling of deference toward the professor or the senior scientist; in fact, some of the European physicists have told me that they were quite taken aback to see Americans and Russians going at each other hammer and tongs in all-out scientific debate at international meetings. The Russians have a very active high-energy-physics program, and are well along with the construction of a seventy-billion-electron-volt accelerator at Serpukhov, which will be the largest in the world when it starts operating. All the physicists I have spoken with at CERN, including Kaftanov, are very eager for increased East-West coöperation, and hope that the existing political thaw will continue to permit it.

ULTIMATELY, the most important process in a scientific laboratory is the process of constant reciprocal education. At CERN, this is facilitated by the layout of the buildings, which are low and long and are joined by a maze of passageways. (The buildings are mostly white with a blue trim, which gives them a clean-cut Swiss look.) As one walks down the halls, one hears a continual buzz of multilingual conversations about physics. There are often knots of physicists in the halls or in the library, which has a few special soundproof rooms with blackboards for informal discussions. Everywhere, one gets the impression of people working and arguing with each other, and this extends even to the cafeteria. There is a long lunch period at CERN (the working day is from eight-thirty to five-thirty, and for many of the experimenters, who work in shifts on the accelerator, it runs into the evenings and weekends), and during it everything closes down—the bank, the post office, the machine shops, and the rest. But the talk goes on. The cafeteria is

furnished with long tables, and by some sort of informal tradition the technical personnel tend to eat at noon, while the physicists eat at one. Usually, the experimental groups eat together and the theorists, too, form groups, sometimes according to language and sometimes according to common interests in physics. After lunch, dessert and coffee are served at a small bar, and everyone spends the rest of the lunch hour in the lounge over coffee or, on sunny days, on the broad terrace in front of the cafeteria, from which one has a fine view of Mont Blanc. Everywhere one looks, there are people discussing physics, sometimes with paper and pencil, sometimes with elaborate gesticulations, and usually in two or three languages. It is the time of day when one hears the latest technical gossip, both from CERN and from laboratories around the world.

In addition to this informal process of education, there are more formal lecture courses and seminars. The summer before last, I attended a lecture series, given especially for physicists, on using electronic computers. Surprisingly, most of the computer use at CERN and at other high-energy-physics laboratories is not by theoretical physicists but by experimenters. A typical experiment involves placing a target, such as a bubble chamber filled with liquid hydrogen or liquid helium, in front of the beam of particles emerging from the accelerator. The particles leave tracks in the liquid, and these tracks are photographed—a process likely to involve photographing hundreds of thousands of tracks from several angles. Then the photographs, which often look like examples of abstract art, must be "scanned;" that is, the events of special interest must be distinguished from the inevitable chaotic background. Much of this scanning is done—visually, in the first instance—by a large group of people, mostly women. The scanners do not have to be physicists, since picking out events of interest is a question of pattern recognition and can be taught to almost anyone. After the events have been roughly selected, they must be "measured." The curvature and thickness of the tracks as well as the angles between them are determined, to see whether the event in question is really what one is looking for or is perhaps something that looks similar but is really quite different. These distinctions are made with the help of a computer, which is programmed to correlate the results of the measurements, try to fit the event with various hypotheses, and then report back. Without a computer, this procedure would be enormously time-consuming, since many possibilities must be explored in each photograph, and there are thousands of photographs to study. Moreover, some devices that make possible a partial automation of the measuring process are now in use—an operator sets a crosshair on a track, and the machine does the rest of the measuring automatically, feeding the results into the computer—and there are systems under development that in certain cases will do the pattern recognition automatically. Hence, one can imagine a time when computers will study all the pictures and deliver carefully analyzed experimental curves to the researcher. The amount of computing required for such work is tremendous. CERN has recently bought the largest computer in the world and will install it at the end of this year, to replace the present equipment, which is completely saturated.

This past summer, I attended two courses given by theoretical physicists especially for the experimenters at the laboratory. There is a communication problem between experimental and theoretical physicists that arises from the increasing need to specialize in a single aspect of physics because of the complexity of the field. The old-fashioned romantic notion of the experimenter coming into the physics laboratory in his white coat, with his mind unburdened by preconceptions or theoretical fancies, and saying to himself, "Well, what am I going to discover today?" just doesn't apply to experimental high-energy physics. The probable theoretical implications of experiments are carefully considered in advance. Recently, in an editorial in *Physical Review Letters*, a journal that specializes in the rapid publication of important new results in physics, Dr. S. A. Goudsmit commented, somewhat ironically,

"At present, most experiments are only undertaken to prove or disprove a theory. In fact, some experimental teams employ a theorist somewhat in the role of a court astrologer, to tell them whether the stars in the theoretical heavens favor the experiments they are planning."

In any case, an experimenter must have a knowledge of the latest theoretical results and how they bear on his work. Thus, one of the jobs of the theoreticians at CERN is to explain what is happening in their fields. One of the special courses, given by Professor Leon Van Hove, a Belgian physicist (formerly of Utrecht, Holland) who directs the theoretical group at CERN, presented an especially lucid review of general aspects of reactions at high energies, but this course was finishing for the summer when I arrived, so I could attend only the last few lectures. The other course, given by Professor Bernard d'Espagnat, a French theorist from Paris, was concerned with some of the most exciting ideas that have come along in elementary-particle physics for several years. These ideas have to do with what is known as "unitary symmetry," or, less accurately, "the eightfold way." To understand what they signify, one must go back into the history of the subject a bit.

In the past few years, more and more new particles have been discovered in experiments with the accelerators. These particles are characterized by, among other properties, their masses, their electric charges, and—because they are in general unstable—their lifetimes. The major problem the particles have presented has been whether they have any interconnections or are completely independent units. In this area, atomic physics furnishes an especially encouraging example, since a superficial look at the array of chemical elements and their diverse properties might lead one to conclude that they could have no connections with one another. However, it is well known that all atoms are composed of only three distinct types of particle—the proton and the neutron, which form the atomic nucleus, and the electron, a light, negatively charged particle that generates a cloud of negative charge around the nucleus. The number and distribution of the electrons determine the chemical properties of a given atom, and the protons and neutrons determine its mass. In the case of the so-called elementary particles, one may ask the same sort of question: Is there a simple basic set of elementary particles from which all the others can be constructed? Or, as the question has sometimes been phrased: Are some elementary particles more elementary than others, and can the rest be made up of the most elementary ones? It is quite possible that this question has no real answer. Observations made with the aid of bubble chambers and other detection devices show that, in accordance with certain general rules, elementary particles can be transformed into one another in high-energy reactions. For example, if a pi-meson from an accelerator bombards a liquid-hydrogen target, there can be reactions in which the pi-meson and the proton that composes the liquid-hydrogen nucleus disappear and out come a so-called K-meson and another particle, called a lambda. Thus, the system of pi-meson and proton is transformed into K-meson and lambda. In accounting for this transformation, one may think of the proton as being made up of a K-meson and a lambda, or one may think of the lambda as being made up of a proton and a K-meson, or one may think of all these particles as elementary. Many physicists have come to believe that the choice among these possibilities is a matter of convenience, to be decided only by which choice leads to the simplest and most beautiful theory. It has recently become clear that all known particles can be thought of as being made up of three basic particles, and this way of looking at them appears to be the simplest possible. The basic set has not yet actually been seen, and one of the great tasks of high-energy experimental physics in the next few years will be to search for new particles that may be candidates for the basic ones. The search has already started at CERN and Brookhaven. The term "eightfold way" derives from the fact that the particles composed of the basic threes fall naturally into groups

of eights (in some cases, into groups of tens) that have closely interconnected properties. There is now very solid evidence that these groupings exist, and if the basic set of threes is identified, this will close one of the most fascinating investigations of elementary-particle physics.

AFTER one of Professor d'Espagnat's lectures, on a particularly warm and lovely summer's day, I decided to take a walking tour of the CERN site. At different times over the years, I had visited most of the installations, but for the fun of it I thought I would make the whole round in one swoop. The laboratory is surrounded by gentle rolling fields leading off to the Jura, the wooded, glacially formed foothills of the Alps; in fact, during the winter, people from CERN often spend their lunch hour skiing in the Jura, which are only a few minutes away by car. When I left the building where the theoreticians have their offices, the first thing that struck me was the construction work going on everywhere—laborers (most of them Spaniards and Italians, as is the case in all of Switzerland) were enlarging roads and erecting new buildings. Alongside one of the roads I saw a striking silvered bubble—a safety tank for holding hydrogen. Hydrogen, which is the most popular target for experiments, because of its simplicity, is also one of the most difficult gases to handle, because of its explosive nature, and there is a whole complex of installations at CERN devoted to processing and handling it, all of them plastered with multilingual signs telling one not to smoke. A little farther on, I came to one of three "halls" in which experiments are actually done. As the proton beam runs around its track, it produces particles in targets, and these can be siphoned off at various stages and directed into one of the halls; this was the East Hall. I am not very enthusiastic about attempts to romanticize science and scientists, but there is something romantic about a high-energy experimental laboratory. Its attraction lies partly in the complexity and diversity of the equipment—giant magnets, trucks filled with liquefied gases, wonderful-looking electronic devices that flash lights of every color—and partly in the knowledge that what is being studied lies at the very heart of the composition of the world. There was almost total silence in the East Hall, broken only by the rhythmic booming of the main magnet of the accelerator and the constant hum of electric motors. (CERN uses almost ten per cent of Geneva's entire power supply.) I stood in awe until someone came up and asked if I was looking for something. For want of anything better, I told him that I had got lost while trying to find the road leading to the center of the accelerator ring. He gave me some directions. I walked outside and quickly found it. The ring is buried, and one can see its outline as a slight circular mound raised above the fields. The center of the ring is guarded by fences and signs warning against radioactivity and barring entry to anyone without permission. This day, though, I noticed a number of men inside the ring cutting the grass; the machine was undoubtedly off while they were working. I crossed over and went belowground into the central building. Inside, equipment sprawled everywhere, and there was a faint smell of resin, which is used in soldering electrical circuits. Dozens of men in laboratory coats were working at one job or another with great concentration. As I watched them, the title of a book on mountain-climbing came to mind—"Les Conquérants de l'Inutile." In a way, high-energy physics is "*la conquête de l'inutile*," but it is also one of the most exciting, benign, and revealing intellectual disciplines that man has been able to devise.

Radioactive materials are being used widely in industry, medical and ecological research, clinical therapy, agriculture, and food processing.

11 The World of New Atoms and of Ionizing Radiations

V. Lawrence Parsegian and others

Sections of a textbook published in 1968.

21.11 *The world of new atoms and of ionizing radiation*

We have gained, as by-products of atomic power, very many new types of radioactive atoms or radioisotopes. There are now about 1100 nuclides that are new and man-made. Each is unstable, but changes in its own time to a more stable form. The change is accompanied by the emission of radiation, either in the form of a γ-ray photon, β-ray, sometimes positron, an α-particle, or some other form or combination. Each nuclide has the chemical properties of a stable, conventional atom, but in addition each also emits radiation of a type and energy that is characteristic of that nuclide. Also, each unstable nuclide (radioisotope) has a particular time rate or half-life for its change of form.

The early forms of Mendeleev's Periodic Table of the atoms listed up to 92 elements. Within the limited science and technology revolving around the chemistry of these elements, there were built up vast chemical industries. The chart of over 1300 nuclides now offers a much larger variety of atoms and building blocks out of which to develop an understanding of atomic behavior.

For example, consider the isotopes of carbon. Two stable forms of carbon are found in nature, one of mass 12 (C^{12}) and one of mass 13 (C^{13}). When nitrogen-14 (N^{14}) is bombarded by neutrons, it captures a neutron and emits a proton, leaving a new atom which has six protons and which therefore behaves chemically like carbon. This is the isotope C^{14}, which is unstable and eventually emits a weak β-particle as it reverts back to the original stable N^{14}. The half-life for this transition is very long, about 5700 years, and the β-ray energy is 0.156 MeV.

These C^{14} atoms become important for several purposes.† They may be incorporated into drugs that contain carbon. When the drug is injected into man or animal (or incorporated into carbon dioxide gas, which may be absorbed by a plant), it becomes possible to follow the course of the carbon in these systems simply by "tracing" the behavior of the C^{14} components; this is done by measuring the radiation they emit. Both time rate and distribution of the drug (or CO_2) in these complex systems can then be determined even though the systems themselves are already full of carbon atoms. This process has made it possible to identify a long series of intermediate steps in the photosynthesis of carbon dioxide for plant growth. The use of radioactive carbon (C^{14}) and radioactive

† We have already discussed the use of C^{14} in radioactivity dating techniques in Chapters 2 and 20.

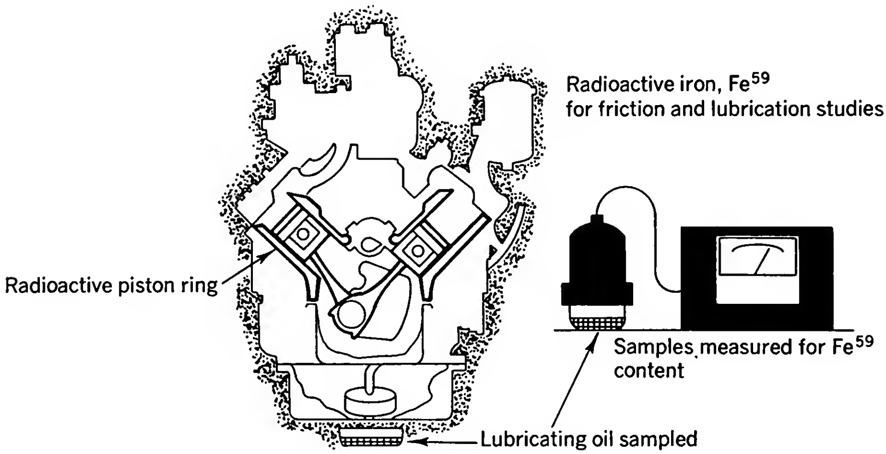


Fig. 21.13. A common application for use of radioisotope iron-59 to measure wear of metal parts. The piston rings are first made radioactive by exposing them to neutrons in a nuclear reactor, then installed in a motor which is under test for wear characteristics. As the piston ring loses metal to the oil, the presence of radioactivity in the oil gives a measure of the wear while the motor is running. When the motor is disassembled, the transfer of metal to the cylinder wall can also be measured accurately. Advantages: (1) transfer of metal measured to $\frac{1}{1,000,000}$ oz.; (2) oil sampled during operation of motor; (3) rapid, simple, economical. (Courtesy of U.S. Atomic Energy Commission.)

species of salts has clarified the understanding of many of the biological processes involved in human blood flow, the diffusion of salts across body membranes, and metabolic activity. Industry has found activation analysis to be particularly sensitive to contaminants in metals or other materials and has used it for identifying these contaminants. Considerable literature has been written about the characteristics and uses of radioisotopes. Many useful publications and references are available through the AEC.

Figures 21.13, 21.14, 21.15, 21.16, and 21.17 illustrate some applications involving radioisotopes.

Radiotracer and dating techniques require relatively weak concentrations of C^{14}, of the order of microcuries. In such applications all that is required of the emitted radiation is that it be measurable, either with Geiger (or similar) counters or with photographic film.

The various types and energies of radiation have penetrating power of differing orders. For example, α-particles can be stopped by a sheet of paper; β-particles may require from several sheets of paper to inches of solid material to stop them, depending on their energy. Gamma rays can penetrate inches of lead. By selecting suitable radiation, one may easily construct gauges for industrial

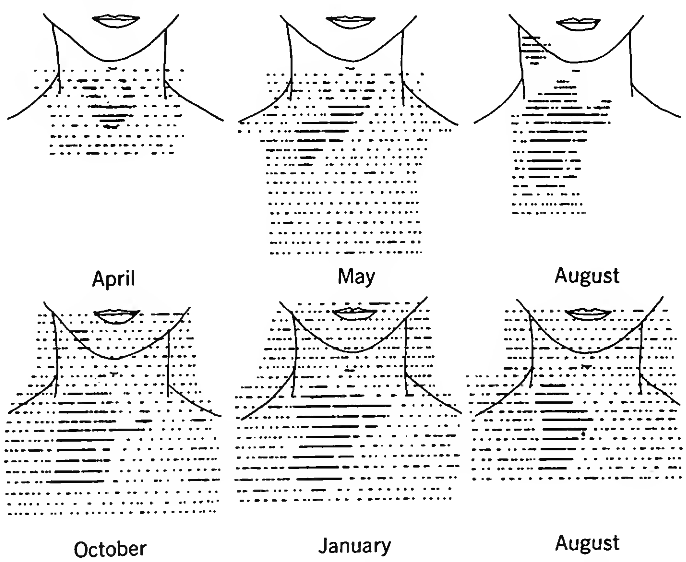
applications that may be used for a wide range of thicknesses.

As noted earlier, the analytic technique called *activation analysis* has become important for industry as well as for research.† If a specimen has a very small amount or trace of impurities and is placed in the neutron flux of a nuclear reactor, the trace impurities (as well as the main body of the specimen in some cases) become radioactive. In many cases the type and amount of the impurity can be determined by comparing the results of irradiation of the unknown sample with the results one obtains by irradiating specimens with known impurities.

The sensitivity of activation analysis is illustrated by the following case: Ordinary arsenic, arsenic-75, on capturing a neutron, becomes radioactive arsenic,

† The term *activation analysis* refers to the process of making a material (which may be a contaminant) radioactive by bombardment with suitable nuclear radiation.

Fig. 21.14. Thyroid cancer. This is a series of six radioiodine scans of the neck and chest of a patient with cancer of the thyroid, made over a period of 16 months at the Oak Ridge Institute of Nuclear Studies. The initial scan (top, left) shows the pattern of the normal thyroid tissue (dark lines) and the presence of the tumor is questionable. With subsequent therapeutic doses of radioiodine, the normal thyroid is progressively fainter and the tumor becomes more apparent as it takes up the radioiodine. Finally, shrinkage in the size of the tumor begins (lower, right scan) as a result of the radioiodine therapy. (AEC Report for 1965.)



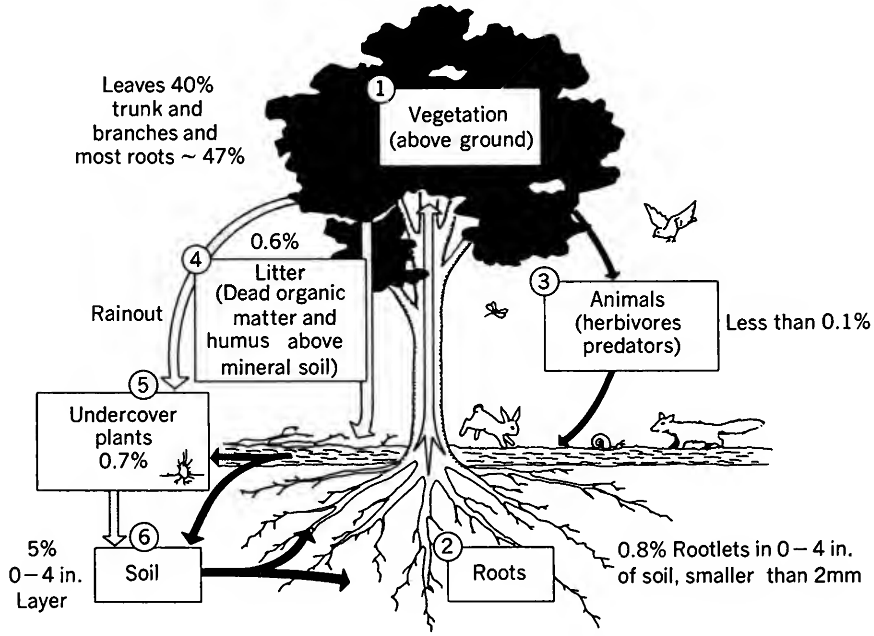
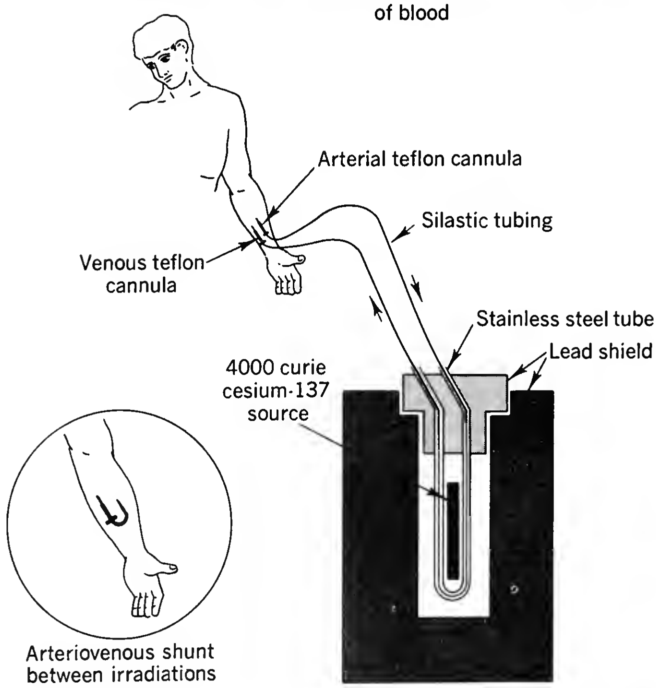


Fig. 21.15. Ecological cycle. The rapid ecological movement of a radioisotope such as a cesium-134 is illustrated in the above drawing of a white oak tree whose trunk was tagged with two millicuries of the radioisotope. Within 165 days, the tracer had become redistributed in different parts of the ecological system and was again entering the tree, this time through the root system. Use of radioisotopes in such studies in the forests at Oak Ridge National Laboratory helps ecologists understand the basic processes that maintain our forest resources. (AEC Report for 1965.)

Fig. 21.16. (Facing page) Treatment of leukemia by irradiation of blood. *A patient at the Medical Research Center at Brookhaven National Laboratory is shown in the photo* top *undergoing treatment for leukemia by extracorporeal irradiation of his blood. The nurse is about to connect the arteriovenous shunt in the patient's forearm to the tubing leading into a shielded container where the gamma-ray source is located. The technique, as diagrammed* below, *was applied to the study and treatment of human leukemia following extensive studies of the origin, function, and turnover rates of cells and other blood constituents of normal and leukemic cows. The purpose of this form of treatment is to destroy leukemic white cells in the blood without injuring other cells or organs in the body; the red blood cells are much more resistant to radiation damage than the leukemic cells. A semipermanent external arteriovenous shunt, which may last for many months, is inserted in the patient's forearm. Arterial blood is propelled by the action of the heart through plastic tubing into the shielded container, past an intense source of gamma rays, and back into the patient's arm. As the blood passes through the gamma source (4000 curies of cesium-137) it receives a radiation dose of from 250 to 900* rads, *depending upon its flow rate* (900 rads *would be a lethal dose of radiation if applied to the whole body). The treatment can be repeated as necessary to reduce the numbers of leukemic cells in the blood. (Courtesy Brookhaven National Laboratory.)*

Schematic diagram of extracorporeal irradiation of blood



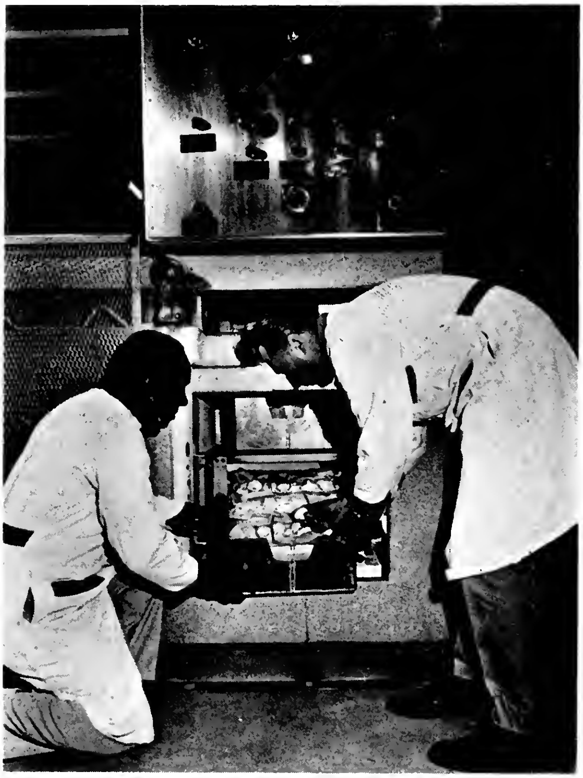


Fig. 21.17. Irradiation of food with ionizing radiation to increase shelf life against spoilage. (*Courtesy Brookhaven National Laboratory.*)

arsenic-76, which emits beta and gamma radiation on decay. Therefore, by radioactive assay, one can determine the concentration of arsenic in a sample. In 1961 a group of Scottish and Swedish scientists subjected a few strands of hair, cut from the head of Napoleon at his death in 1821, to neutron irradiation and found arsenic to be present in thirteen times normal concentration, thus suggesting that Napoleon might have been poisoned. Closer investigation indicated a definite pattern of the variation of arsenic concentration in the hair. This pattern, when compared with the record of Napoleon's sickness, revealed a correlation with his periods of severest pain. It seems arsenic was in the medicine given to relieve his pain and it may have had untoward effects as well.

21.12 Effects and products of ionizing radiation

The ionizing radiation given off by radioactive isotopes can be concentrated and intense. Since this radiation is highly penetrating and ionizing, and induces changes in biological and chemical systems, it promises to become significant in chemical processing and in destroying unwanted bacteria (such as in milk) and tissues (such as in tumors, cancers). But this promise is a mixed blessing and curse, for overexposure to radiation is a health hazard. It has been found to cause leucopenia (decrease in number of white cells in blood), epilation (loss of hair), sterility, cancer, mutations (altered heredity of offspring), bone necrosis (destruction and death of bones), and eye cataracts.

In conventional processes, chemical reactions proceed as a result of atomic collision, favorable valence combinations, excitement of atom systems by heating, Coulomb attraction, free radical intermediates, and other similar activators. The energy exchanges are likely to be of

the order of a few electron volts or less per atom (or molecule).

When swift, charged particles (such as α-particles, protons, or β-particles) pass through matter, they leave tracks of ionized and excited atoms and molecules, which undergo vigorous reorganization. The concentration of energy can be hundreds or more times the intensity of conventional processes, especially with heavy charged particles and toward the end of particle tracks in the material. As a result, radiation effects are often deleterious to the properties of the material.

There are, however, applications wherein the destructiveness of radiation is desirable, such as for killing insects that infest grain or microbe systems in medical supplies. There are also cases where the reorganization of atoms and molecules following irradiation results in improved physical properties or produces desired chemical changes. Radiation induces such widely different reactions that it becomes a very versatile research tool. Processing by irradiation also appears to have very real possibilities of competing with some conventional industrial processes and of inducing reactions that cannot be produced by other means.

The activities involving radiation and radiation chemistry may be grouped under six categories: food preservation, sterilization, chemical processing, radiography and medical therapy, radioisotope power sources, miscellaneous.

Since ionizing radiation can be lethal to living organisms and microorganisms, one of the early programs sought to sterilize foodstuffs and thus give them longer shelf life. Early efforts concentrated on sterilizing meats and other foods by radiation dosage ranging from 2 to 5 megarads (million rads†). The results of these early years were not successful because the heavy dose caused changes in the taste and appearance of foods. More recent work has been much more encouraging. In 1963 the Food and Drug Administration (FDA) approved sterilization of bacon by gamma radiation (up to 2.2-MeV energy) and by electron beams (up to 5-MeV energy) from accelerators. The sterilization of ham, chicken, and beef appears promising.

When the radiation dose is kept well below the doses required for sterilization, down to values of 500,000 rads or less, the effect is to "pasteurize" foods in a way that often permits longer shelf life. For example, a dose of 250,000 rads will extend the shelf life of haddock fillets to 21 days at 32° to 33°F. Crabmeat treated with 200,000 rads had its shelf life increased from 7 days to 35 days when held at 33°F. Fruits (strawberries, cherries, citrus, pears, tomatoes) show similar gain. Insects in grains and wormy (helminthic) parasites such as those associated with trichinosis from pork are killed by 30,000 rads. Sprouting of potato tubers can be inhibited with doses from 10,000 to 15,000 rads. But dosages in excess of 10 million rads appear to be needed to inactivate some enzymes.

Radiation does not raise the temperature of the processed materials at these dosages. Furthermore, with γ-radiation, the whole process can be mechanized and the foods can be irradiated in the packaged state. The main difficulty is the cost of the radiation, whether one uses radioisotopic sources or an accelerator. The irradiation of fish adds from 1 to 3 cents per pound, which is probably acceptable. Because strawberries may cost about 50 cents per pound, they can stand an irradiation expense of an additional $1\frac{3}{4}$ cents per pound. But for other fruits

† A rad represents the absorption of 100 ergs of radiation energy per gram of absorbing material.

and for grains, the cost probably must remain at $\frac{1}{4}$ cent per pound, to be economically acceptable. To help this industrial development, the Commission has reduced the selling price for certain radioisotopes such as cobalt-60 (Co^{60}, which emits strong γ-rays of 1.1 and 1.3 MeV and has a half-life of 5.3 yr) and cesium-137 (Cs^{137}, which emits gammas of 0.66 MeV with a half-life of about 33 yr).

Radiation costs come down sharply as the radiation intensity of the facility is increased, in terms of kilowatt capacity, for either radioisotopic sources or accelerator sources. But it is difficult to find many geographic sites where one can provide high enough production quantities to bring the cost of radiation pasteurization down to 1 cent per pound.

How about irradiation to sterilize materials that are not foodstuffs, such as medical supplies, sutures, bandages, and drugs? While there are limitations in this area also, there are some real advantages to radiation processes as compared with the use of heat, chemicals, or ultraviolet light. When penetrating radiation is used, sutures or other supplies can be packaged in conventional work areas and then irradiated while in sealed state.

21.13 Radiography and medical therapy

These two subjects may be treated together because they depend on similar sources and techniques. Gamma rays are very penetrating—more so than X-rays from conventional machines. A cobalt-60 source can therefore be used effectively for penetrating metal parts, castings, tank walls, and the human body. As in X-rays, the radiation that passes through the target or body can be recorded on photographic film or on a fluorescent screen, to give a faithful picture of the variations of matter through which it passes. Flaws, cracks, cavities will show up as clearly as with X-rays.

The advantages of radioisotopic gamma sources over X-ray machines are three:

(1) These sources can be made portable and do not require electric power for their operation.

(2) Radioisotopes emit radiation in all directions, which makes it possible to obtain radiographs all around a vessel into which the source is placed.

(3) Radioisotopes can provide higher penetrating power without requiring excessively large installations.

Very many industrial firms make use of such radioisotopes as Co^{60}, which is equivalent to 2.5-MeV X-rays and can be used for steel of 2- to 5-in. thickness. For lesser penetrability, iridium-192 (Ir^{192}), cesium-134 (Cs^{134}), and Cs^{137} are the equivalent of up to 1400-keV X-rays and are useful for radiographing steel plates from $\frac{1}{2}$- to $2\frac{1}{2}$-in. thickness (or an equivalent density of other materials). Thulium-170 (Tm^{170}), europium-155 (Eu^{155}), and certain isotopes of americium (Am) provide still lower penetration.

For many fixed installations, X-ray machines may be preferred. Some industrial firms engaged in the production and testing of tanks, ships, and transmission pipe in the field have found the radioisotopic sources to be much more practical than X-ray machines.

We have noted that radiation kills living organisms. Malignant disease in body tissue can often be arrested by exposure to penetrating, ionizing radiation. But since healthy tissue also suffers, radiation must be applied carefully and restrictively to the tissues to be treated. This has given rise to very many designs that use radioisotopic sources in the form of tiny needles that are inserted into tissue; or the sources may be contained in a housing that directs a well-collimated beam onto the tissue.

Radioisotopic sources offer portability and considerable choice in the type and energy of radiation that they emit. Also they can be fabricated into very many shapes and sizes.

Different approaches to the nucleus suggest different models. This paper considers several nuclear models including the liquid-drop model, the shell model and the optical model.

12 The Atomic Nucleus

Rudolf E. Peierls

Scientific American article published in 1959.

Ever since 1930, when the discovery of the neutron made it plain that the nuclei of atoms were built of protons and neutrons, physicists have been trying to form a picture of the structure of the nucleus. The same task for the rest of the atom was completed in the first quarter of this century. We were able to understand in detail how the electrons move under the attraction of the nucleus, and how their motion is influenced by their mutual repulsion.

To achieve such an understanding requires three major steps: First, we must know the forces between the particles. Second, we need to know the mechanical laws which govern their motion under the influence of these forces. Third, we need in most cases a simplified picture, or model, from which to start. Once we have the first two ingredients, we could in principle write down a set of mathematical equations whose solutions would tell us all about the atom, or about the nucleus. In the simplest possible atoms, like that of hydrogen, in which there is only one electron, or in the simplest compound nuclei, like the deuteron, which contains only one proton and one neutron, such equations can be written down and solved without difficulty. However, for more complicated structures this head-on attack becomes much harder and soon exceeds the capacity even of modern electronic computers.

We are like men who encounter for the first time a complicated machine, and who try to analyze its operation. If we attempt, without any guidance, to puzzle out the interplay of all the parts of the machine, we should soon lose ourselves in a maze. Instead, we first try to ascertain the major features of the machine's operation. We then devise a model which resembles the real thing in these features, yet is simple enough to be analyzed. Then, of course, we must put in corrections for the complications which we have left out and check that they do not materially alter the picture.

In the study of the atom the first of the three steps hardly presented a problem. As soon as Ernest Rutherford had demonstrated that the atom consisted of a heavy, positively charged nucleus and of light, negatively charged electrons, it was taken for granted that the forces between them were the electric attraction of unlike charges, following the inverse-square law familiar to every student of physics. The major difficulty was the second step. It turned out that the basic mechanical principles of Isaac Newton, which applied to all "large" objects from the planets and the moon down to steam engines and watches, had to be revised in the atomic domain. To understand atoms we had to use the new ideas of the quantum theory, following the pioneer work of Niels Bohr, who adapted for this purpose the concept of the quantum of action which Max Planck had first found in the behavior of light. These new laws of mechanics were later formulated as the laws of "quantum mechanics," or "wave mechanics," which gave us complete command over the theory of the atom.

The third step, of finding a simplified model for discussing the atom, also proved relatively easy. In working out the possible orbits of a single electron under the attraction of a proton, as in the hydrogen atom, Bohr found that one could account for the behavior of a more complex atom by assuming that each of its electrons moved in such an orbit. The larger the number of electrons in an atom, however, the more distinct orbits they occupy; this is a consequence of the "exclusion principle" discovered by Wolfgang Pauli, which limits the number of electrons that can travel in a given orbit.

We must allow not only for the attraction of the electrons by the nucleus, but also for the repulsion of the electrons by one another. However, we simplify the nature of this repulsion by forgetting that it changes continuously as the electrons move around in their orbits, and treating it as a fixed field of force. In other words, we replace the repulsion due to a moving electron by that which we would obtain if the electron were spread out evenly over its orbit. This simplification can be justified by the fact that the repulsion acts over relatively long distances, so that each electron is at any time under the influence of several others. If we underestimate the effect of one of the electrons which may happen to be rather close to the one we are looking at, we are likely to overestimate the effect of another which happens to be rather far away.

This model of the atom is usually called the "shell model," because it is convenient to group together the electrons moving in orbits of similar size but of different shape and direction. Such a group of orbits is called a shell.

When the atomic nucleus first became an object of serious study, the nature of the difficulties was rather different. The general laws of dynamics did not seem to require further revision; the laws of quantum mechanics which had been discovered in atomic physics seemed quite adequate for the nuclear domain. Indeed, we have not yet found any evidence in the behavior of nuclei which would suggest that these laws might be in error. Thus the second step in our list presented no problem.

The Nuclear Forces

On the other hand the first step—the determination of the forces between the particles—proved to be a very difficult

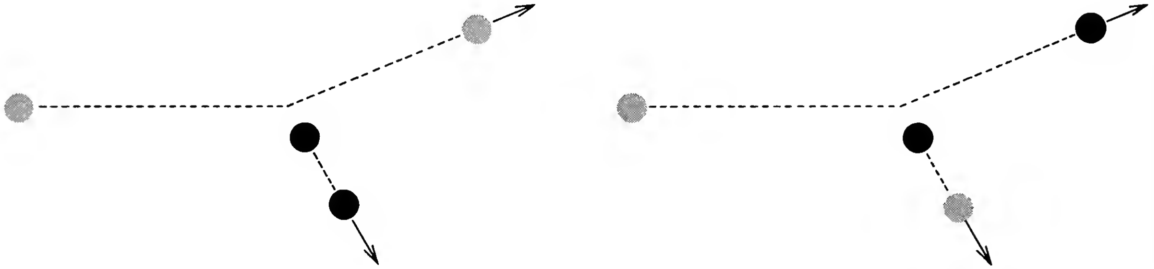
problem. Even today, after some 25 years of intense study, we cannot claim to have a complete answer, but we have by now at least a fair knowledge of what the forces are like.

They cannot be electric in origin. The only electric charges found in the nucleus are the positive charges of the protons, and like charges repel each other; thus electric forces cannot be responsible for holding a nucleus together. Moreover, electric forces are much too weak. We know that the energy of attraction of two unlike charges (*i.e.*, the work we have to do to pull them apart) varies inversely as their distance. The attractive energy of an electron and a proton in the hydrogen atom is a few electron volts (ev), and since the diameter of the hydrogen atom is 20,000 times larger than that of the smallest nucleus we should expect electric energies in the nucleus to amount to some tens of thousands of electron volts. Actually the forces inside a nucleus run to many million electron volts (mev). It follows that nuclear forces are vastly stronger than electric forces.

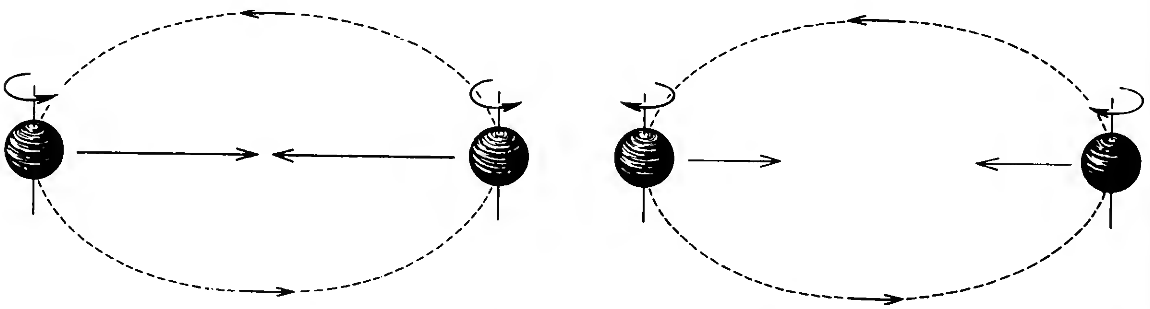
It is also clear that these strong forces act only over extremely short distances. The pioneer work of Rutherford on the passage of charged particles through matter showed that, even in encounters in which a charged particle approaches a nucleus to a distance of a few times the nuclear diameter, the only noticeable force is the electric one. We know today that nuclear forces between two particles are quite negligible if the distance between the particles is more than, say, four fermis. (The fermi, named for the late Enrico Fermi, is a convenient unit of distance for the nucleus. The diameter of a heavy nucleus is some 15 fermis; the diameter of the hydrogen atom, about 100,000 fermis.) It is not surprising, therefore, that earlier physicists did not meet nuclear forces in laboratory experiments. The only possible way of studying these forces is to observe the behavior of nuclei, or to bombard hydrogen or other nuclei with fast protons or neutrons under circumstances in which the effect of really close encounters can show up.

What makes this task harder is that the nature of nuclear forces, unlike the simple inverse-square law of electric or gravitational forces, is rather complicated. If the law of nuclear forces were simple, a few observations might suffice to guess its general form. But all simple guesses based on a few experiments have been disproved by later experiments. We are obliged to reconstruct the law of nuclear forces laboriously from the various pieces of evidence we can extract from the experiments.

Ultimately we hope to be able to derive the law of the forces from more basic principles, just as we can derive the inverse-square law of electric forces from the basic laws of electromagnetism. A beginning was made by the Japanese physicist Hideki Yukawa, who used the analogy with electromagnetic radiation to point out that nuclear forces must be related to a new form of radiation which could carry charged particles weighing a few hundred times more than the electron. His prediction was confirmed by the discovery of the so-called pi meson. His picture of the mechanism underlying the nuclear forces has been qualitatively confirmed by many observations, and has been a useful guide in our thinking about the forces. But it has not yet been possible to use his idea for a reliable and accurate derivation of the law of the forces because of the mathematical problems which stand in the way. We do not know today whether a correct solution of the equations em-



CHARGE EXCHANGE in the nucleus is schematically depicted. When protons (*black balls*) are struck by fast neutrons in half the cases (*left*) the neutron continues forward. In the other half (*right*), the proton exchanges its charge with the neutron.



SPIN-ORBIT FORCE arises from a relationship between spin and orbit. When two particles (*left*) spin in the same direction as that in which they move on an orbit, the force between them is strong. When they spin in opposite directions (*right*), force is weak.

bodying Yukawa's idea would yield the right forces, or whether there is something basically wrong with this approach. The difficulties arise chiefly from the greater strength of the nuclear forces, as compared to electric forces, which makes their mathematical analysis much more difficult.

Thus the best source of information about the forces still lies in direct experiments. These require collisions at high energies—much higher than the energies of particles inside ordinary nuclei. The reason for this is the wave aspect of particles, which is an essential feature of quantum mechanics. Slow particles are associated with waves of long wavelength, and collisions involving such slow particles do not provide much information about the finer features of the forces at work between them, just as in looking through a microscope at a dust particle with a diameter less than a wavelength of light we see only a general blur which does not reveal the shape or nature of the particle. To have particles of sufficiently short wavelength one must raise their energy to a few hundred mev. The most reliable information on nuclear forces has therefore become available only in the last few years, as a consequence of the development of accelerating machines which produce clean beams of protons, neutrons, or electrons with such energies. This need for high-energy beams is entirely similar to the situation in atomic physics, where detailed pictures of the structure of atoms require the use of X-ray or electron beams of several thousand ev—much greater than the energies of the electrons inside the atoms, whose wavelength is comparable to the atomic diameter. The complexity of the results has also made it necessary to call on the services of fast electronic computers for disentangling the observations.

I shall not attempt in this article to give anything like a complete specification of the nuclear forces, but shall stress only those features which are of importance for what follows. We have already noted that the forces must be strong and of short range. Since they hold the different particles together, they must on balance be attractive. At the same time they cannot be entirely attractive, since otherwise heavy nuclei would "collapse." By collapse we mean a state of affairs in which all the particles in a nucleus are so close together that each one is within the range of the attractive force of every other. In that case the attractive energy acting on each particle would grow with the total number of particles present, and the volume occupied by the whole nucleus would be the same no matter how many particles were in it. This is not found in reality. The energy per particle is roughly the same for all nuclei, light or heavy, and the volume of nuclei increases with the number of particles in them.

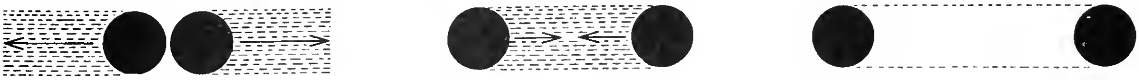
The Exchange Forces

This behavior, which indicates a limited attraction, is usually called "saturation" of the nuclear forces. There are two particularly plausible ideas to account for this saturation. One was suggested by the German physicist Werner Heisenberg, who was one of the founders of quantum mechanics. He postulated that at least part of the nuclear forces between a neutron and a proton involves an exchange of their position, so that after an encounter between them the neutron would tend to follow what had been the path of the proton, and vice versa. The exchange occurs readily only if the two move in very similar orbits, and, since the Pauli exclusion principle allows only a limited number of particles to follow the same orbit, such exchange forces would expose each particle to a strong attraction only from a few others. The bombardment of protons with fast neutrons confirmed this idea, because it showed that in most cases either the neutron or the proton tended to go forward with almost the same speed and direction with which the neutron had arrived. Since it is hard to deflect such fast particles from their path, this indicates that the incident neutron had continued almost in a straight line, but that in half the collisions it had changed its nature and become a proton, leaving a neutron behind.

However, the experiment also showed that only one half of the force was of the exchange type; the other half (corresponding to the neutrons still moving forward after collision) was an "ordinary" force. This is not enough to yield the required saturation, and some other factor must be involved. The second factor tending toward saturation is almost certainly a reversal of the direction of the nuclear forces at short distances, so that, as two particles approach each other, the attraction changes to repulsion. This concept of "repulsive cores" in the forces is familiar in the behavior of atoms. When atoms form chemical compounds, or liquid or solid substances, they are held together by attractive forces; but each atom has a fairly definite size, and when two atoms come into actual contact, their attraction changes into repulsion. We may liken this behavior to that of two rubber balls tied together with a rubber band. There is an attraction between the balls, but there is also a contact force which prevents the centers of the balls from approaching each other closer than one diameter. Shortly after the theoretical need for such a repulsive core in the nuclear forces had become clear, experiments on collisions between fast particles indeed showed direct evidence for these repulsive forces.

Among other features of the nucleus I should mention the "spin-orbit" force, that is, the dependence of the mutual interaction of two particles upon the direction of their orbit with respect to their spin. When the two particles spin on their axes in the same direction as that in which they revolve about each other, the attraction between them is stronger; when they spin in the opposite direction from that in which they revolve, the attraction is weaker. There is some evidence for such a spin-orbit force in experiments on nuclear collisions, but there is still some room for controversy in the interpretation of these experiments.

Our present knowledge of the nuclear forces, while still incomplete, is sufficient to discuss the behavior of nuclei and the collisions between them. At this point we meet the need for the third step in our general program, namely a simple model in terms of which we may approach the dynamical problem of the



NUCLEAR FORCES are dependent on the distance between particles. If the particles are very close, they repel each other (*left*). If they are a certain distance apart, they attract each other (*center*). If they are farther apart, they have little effect on each other (*right*).

motion of the 16 particles in the oxygen nucleus, or the 208 particles in the most stable lead nucleus.

Models of the Nucleus

The selection of a suitable model is not at all straightforward. Not that there is a shortage of suggestions. In fact the trouble in the recent past has been a surfeit of different models, each of them successful in explaining the behavior of nuclei in some situations, and each in apparent contradiction with other successful models or with our ideas about nuclear forces. In the past few years great progress has been made in bringing some order into this confusion and in understanding the justification for each of the models in the domain to which it is properly applied. I shall attempt to explain briefly some of the ideas behind these developments.

The most obvious idea was to use the shell model, which had been so successful in dealing with the atom. In fact, the first attempts to set up such a shell model were made even before the discovery of the neutron, when it was believed that nuclei were made of protons and electrons. A shell model with the wrong constituents cannot have much success in accounting for the facts, but in those days rather few facts were known, so such models were able to survive for some time.

After the discovery of the neutron, attempts to formulate a nuclear shell-model were renewed. This involved the idea of orbits (or quantum states) for the protons and neutrons, in which each of them was pictured as moving independently under the influence of some force which represented the average effect of the others, as in the case of the electrons in the atom. It did not seem possible, however, to choose groups of orbits of the right kind, so that the number of similar orbits which formed a shell could accommodate just the right number of neutrons and protons to account for the exceptional stability of nuclei with certain numbers ("magic numbers") of neutrons or protons.

The same idea was applied to the collision of neutrons with nuclei. According to the shell model, the impinging neutron should travel through the nucleus on its own orbit, as through some field of force, and individual encounters with the particles constituting the target nucleus ought to be rare and unimportant. Hence the neutron should in most cases emerge with the same speed as that with which it entered, and only rarely should it get trapped. The details of the process should not depend critically on the speed of the neutron.

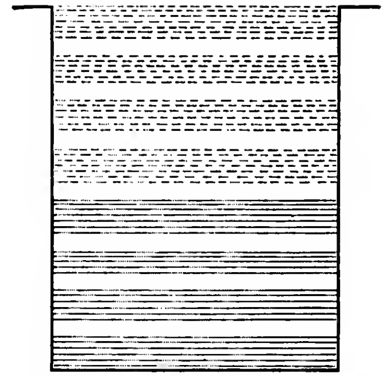
Observations of such collisions, initiated by Fermi in Rome, gave a completely different picture. Most of the neutrons that interacted with a nucleus were trapped, their excess energy being radiated in the form of gamma rays. Moreover, the chance of the neutron being affected by the nucleus depended very critically on its energy. One found a large number of resonances, *i.e.*, sharply selected energies, for which a neutron was sure to be picked up by the nucleus. For each target nucleus there are many such resonances, the energy difference between them being often as low as 100 ev, an exceedingly small difference on the nuclear scale.

These resonances turned out to be exceedingly sharp, and on the uncertainty principle of quantum mechanics a sharply defined energy is associated with a long time. So it follows that once a neutron gets into a nucleus in conditions of resonance it must stay there a long time —much longer than it would take it to cross a region the size of a nucleus.

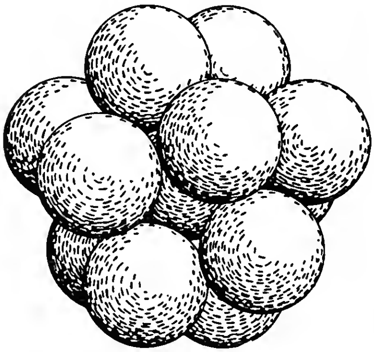
The Liquid-Drop Model

The way to resolve these apparent contradictions was pointed out by Bohr. He recognized that it was not right to think of a neutron as passing just through a general field of force, since the nucleus is densely packed with particles which each exert strong forces on the extra neutron as well as on each other. Instead of comparing the process with the passage of a comet through the solar system, as was appropriate for the passage of an electron through an atom, we should liken it to the entry of a golf ball into a space already fairly densely filled with similar balls. The result will be a complicated motion of all the balls, and the energy of motion of the extra one will rapidly get shared with the others.

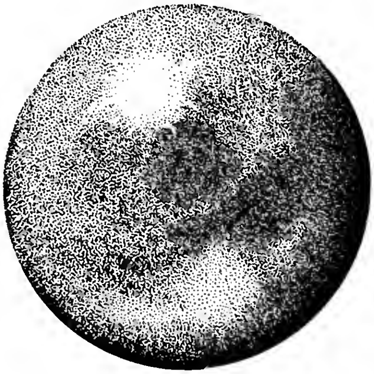
The dynamical problem is now that of a true many-body motion, and we have vastly more possibilities of varying the details of the motion of all the particles. This means that the rules of quantum mechanics will give us far more states of motion, and these are responsible for the greatly increased number of resonances. We also see the reason for the long stay of the neutron in the nucleus, because when the energy of motion is shared among many particles, none of them can attain enough speed to escape from the general attraction. It must take a long time before by chance one of them col-



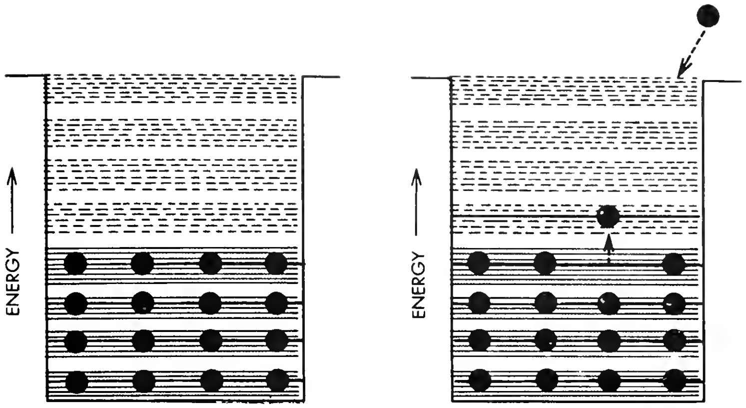
SHELL MODEL of the nucleus is represented by a potential "well" in which the groups of horizontal lines indicate orbits that can be occupied by particles in the nucleus. The groups of solid gray lines indicate orbits of lower energy; the groups of broken gray lines represent orbits of higher energy.



LIQUID-DROP MODEL may also be represented as a collection of golf balls. When another particle, or golf ball, enters the nucleus, the motion of all the balls is disturbed.



OPTICAL MODEL pictures the nucleus as a somewhat cloudy crystal ball. The cloudiness represents the tendency of bombarding neutrons to be absorbed by the nucleus.



LOW-ENERGY ORBITS in the shell model of the nucleus may each be occupied by only two neutrons (*colored balls*) and two protons (*black balls*). In the normal state of affairs (*left*) the low-energy orbits are filled; the particles cannot gain or lose energy, and thus cannot change their orbits. A bombarding particle (*upper right*) has energy to spare; thus it can exchange energy with a particle in nucleus and move it to orbit of higher energy.

lects enough of the available energy to get away. In our picture of the golf balls this will actually never happen, because in the meantime too much of the energy will have been dissipated in friction. In the nuclear case the analogue of friction is the loss of energy by gamma radiation, and this is responsible for the events in which the neutron gets trapped. But it is less effective than in the case of the golf balls, and some neutrons do get out again.

The physicist does not invoke here the similarity with a system of golf balls, which is not quite close enough, but he is reminded of a very similar situation which arises when a water molecule hits a drop of water, and for this reason Bohr's model is often called the "liquid-drop model."

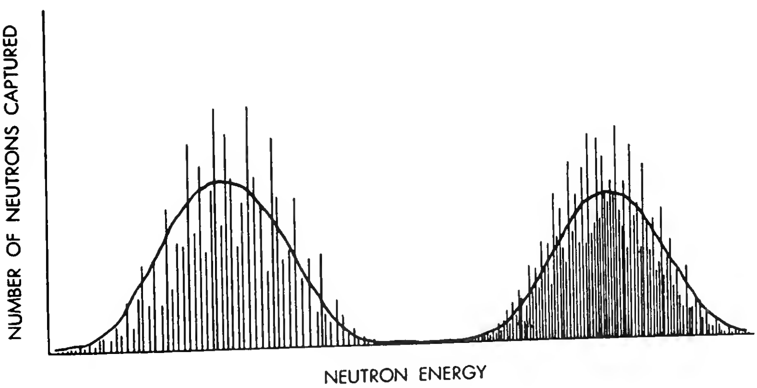
The liquid-drop model met with considerable success, and was able to explain many detailed features of nuclear reactions. At this time it seemed evident that the whole earlier idea of the shell model, which pictures the particles as moving independently, was doomed to failure, in view of the high density of the nucleus and the strong forces a particle was bound to experience in many encounters with others during the course of its motion. Most physicists then regarded the whole idea of a shell model as misconceived, but some, whether out of a stubborn refusal to accept the arguments against the model, or out of a deeper intuitive insight which convinced them that somehow one might be able to get around the argument, continued to look at the behavior of nuclei in their normal states in terms of shells.

The Shell Model Again

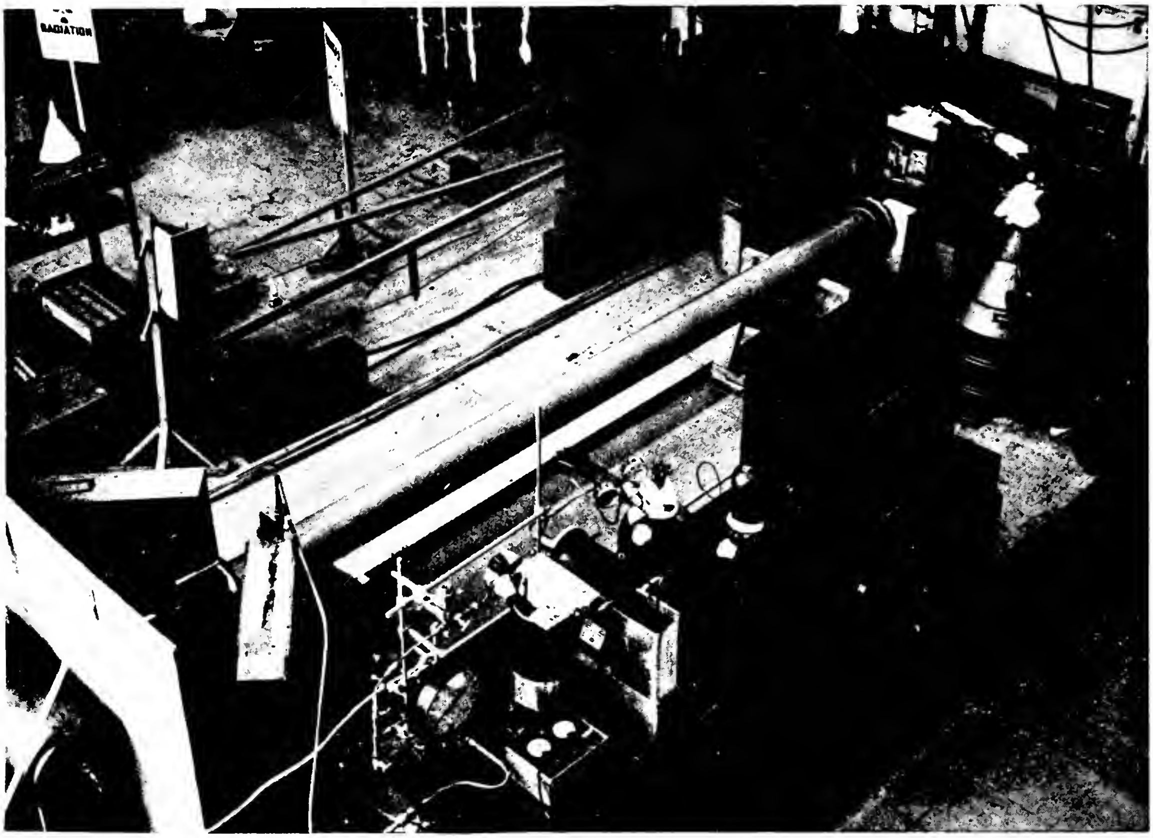
It soon became evident that there was overwhelming evidence in favor of such a shell picture, and the final success came when Maria G. Mayer of the University of Chicago and J. D. H. Jensen of Heidelberg independently noticed that the facts fitted amazingly well with a slightly modified shell model. The new feature was that when a particle spins in the direction in which it moves about the center of the nucleus, its orbit is different from the orbit of a particle spinning in the opposite direction. When this idea was put forward, it was not known that the force between two particles depends on the relative orientation of spin and orbit. Today the idea appears entirely natural. With this refinement, such a mass of data about the behavior of nuclei could be explained that there remained no doubt as to the essential of the particle being absorbed, *i.e.*, lost from the beam of bombarding neutrons [see "A Model of the Nucleus," by Victor F. Weisskopf and E. P. Rosenbaum; SCIENTIFIC AMERICAN, December, 1955]. How can we understand the success of this picture of independent particle motion in view of the Bohr argument?

The answer to this question has been given in essence by Weisskopf. It may be expressed by considering the time sequence of events. To be sure, the bombarding particle is likely to be disturbed from its path by collisions, but this will take a little time. So for a short time it will penetrate into the nucleus on a regular orbit, and this initial period is important for determining whether it will actually get deep inside or be turned back at the surface. Now, to recall once again the uncertainty principle, we know that in talking about a short time interval we must not try to specify the energy too accurately. We should therefore think not of neutrons with a well-defined energy, but of a beam of neutrons varying in energy by an amount that is greater the shorter the time in which they are likely to be involved in collisions inside the nucleus. Experiments often make use of such mixed beams, if the experimenter does not take trouble to select the neutron energies accurately. If we have data with accurate energy selection we should lump together the observations over a suitable range of energies.

Then we do not see the sharp resonances any more because there will al-



GIANT RESONANCES of a typical nucleus are indicated by the colored curve. Each of the vertical lines represents an ordinary resonance. The height of each line denotes the number of bombarding neutrons at that energy which are trapped within the nucleus, or which emerge from the nucleus with only part of their original energy. Giant resonances are observed when nucleus is bombarded with particles of lower energy and lower resolution.



OXYGEN NUCLEI ARE BOMBARDED with neutrons in this apparatus at the Brookhaven National Laboratory. The neutrons are produced by the Brookhaven nuclear reactor, the concrete shield of which is visible at right. The oxygen atoms are contained in the long tank in the middle of the picture. The neutrons which are not absorbed are counted in the shorter tank at lower left.

ways be many of them within the energy range we use. The result we get in this way will reflect the number and strength of the resonances within the selected range. But we may now think of these results also as determined by the first short time interval of the event, and as the neutron pursues a regular orbit during this short time interval the results now should reflect the behavior of such regular orbits. This therefore leads us directly to the picture of the optical model, which has neutrons traveling in regular orbits. The absorption which was allowed for in Weisskopf's optical model merely reflects the fact that the particles do not stay on such a regular orbit forever, but are sooner or later removed from it by collisions with other particles.

The strength of this absorption is thus related to the rate at which collisions occur inside the nucleus. If they are very frequent, so that the particle covers only a small fraction of the nuclear diameter before it hits something, the "giant resonances," which correspond to the orbits of a single particle, will become weaker and more diffuse. The fact that they are found to be pronounced and distinct shows that the particle has a fair chance of completing at least one revolution in its orbit. In this respect we see that the extreme form of Bohr's liquid-drop model, or our simple picture of golf balls, exaggerates the situation. But we have succeeded in reconciling Bohr's explanation of the many sharp resonances in terms of the many-body aspects of the problem, with the superimposed structure of giant resonances, which characterize the early stages of the process.

It remains to account for the quantitative features of the optical model—and in particular for the long time a particle can stay in its orbit before being thrown out of it by a close encounter with another particle—in terms of the basic forces. A promising attack on this problem is now under way. The workers engaged in it include G. E. Brown in the author's group at the University of Birmingham. In particular, the low rate of collisions is seen to be linked again with the effect of the exclusion principle. We have seen that this cuts down the rate of collisions in a normal nucleus drastically. In the impact problems where there is more energy to spare, the collisions are more frequent, because there are more orbits available that are not already occupied, but the prohibition is still partly effective and the collision rate is still a good deal less than that suggested by the picture of golf balls, for which all quantum effects, including the exclusion principle, are of no importance.

A picture thus emerges in which the various, apparently contradictory, models of the nucleus are seen as consistent parts of a whole, each appropriate for answering certain questions about the behavior of nuclei. There are problems for which yet other models have to be used, including the important "collective model" developed by Aage Bohr and B. Mottelson of Copenhagen, but it would exceed the scope of this article to describe them and show how they fit into the story.

The origin of the sun's energy is a long-standing scientific problem. The answer came eventually not from astronomical studies alone, but from investigations of the behavior of elementary particles.

13 Power from the Stars

Ralph E. Lapp

Chapter from his book, *Roads to Discovery*, published in 1960.

THE billions upon billions of stars in the vast universe all have one thing in common—they are all immense masses of flaming gas. Heat evolved deep within this fiery sphere gives rise to the brilliant light which makes the star visible. Our nearest star—our sun—is the source of life on earth. Our planet is kept warm, the oceans remain unfrozen and crops grow because of solar warmth.

Our planet, earth, is but a small sphere some eight thousand miles from rim to rim. It whirls through space and, caught in the invisible grip of the sun's gravitational attraction, orbits endlessly, maintaining an average distance from the sun of 93 million miles. At this distance the earth receives only a minute fraction of the vast outpouring of heat and light that the sun radiates. In fact, two billion times more heat flies off into space than strikes the earth.

How does our sun manage to keep its heat furnaces stoked? How has it kept blazing away at this rate for five billion years? Is there any danger that it may "run out of gas"?

Only recently, with the data turned up in nuclear research, has it been possible to answer these questions. Yet from the time of the primitive caveman, the sun has been an object of

wonder and of worship. The ancients revered the Sun God and countless humans were sacrificed on bloody altars to assuage the fiery deity.

In more modern times wonder turned to curiosity and curiosity to methodical investigation. Astronomers found that the sun is a million times bigger than the earth, that the temperature at the sun's surface is about six thousand degrees Centigrade, and that the temperature deep inside the core must be about fifteen million degrees Centigrade. Astrophysicists proved that no ordinary burning or chemical combustion could account for solar heat. They knew there was not enough oxygen to support such a combustion. All efforts to explain the sun's power failed; no energy source was powerful enough to account for such flaming heat over a period of five billion years. By all reckoning, the sun should have spent its energy long ago; it should be a dead cinder in the sky surrounded by lifeless, frozen planets—a darkness in the universe.

Sir Arthur Eddington was the first scientist to speculate correctly about the source of the sun's energy. He suggested in 1920 that stars might gain energy from the combination or fusion of hydrogen to form more complex elements. This nuclear "burning" should release per atom a million times more energy than any known chemical process. Eight years later Frederic Houtermans and Robert Atkinson took the next step which turned speculation into theory. They calculated that hydrogen within the sun's core consisted of atoms so speedy (due to heat and pressure) that some collisions between hydrogen atoms would produce a thermonuclear reaction with the release of heat. We call this thermonuclear energy and, as the name implies, it is nuclear energy produced by heat-agitated atoms.

Houtermans and Atkinson had practically no experimental data about the behavior of hydrogen atoms, so they had to proceed on pure theory. They knew that at the elevated temperatures inside the sun's core hydrogen atoms would be

stripped of their electrons. They also knew that the great pressure due to the overweight of the sun's voluminous mass squeezed hydrogen nuclei (protons) so close together that the result was a proton paste eight times denser than solid lead. Houtermans and Atkinson calculated that hydrogen fusion could account for solar heat. However, they could not demonstrate that the fiery proton paste in the sun's core would actually sustain a thermonuclear reaction. They lacked the vital nuclear data to predict the behavior of protons at the temperature that exists inside our sun.

At this point we must pause to show that the "temperature" and "energy" of protons or, for that matter, any particle, may be related. This is important because the nuclear behavior of a particle depends very strongly upon its energy (or its speed).

Ordinarily, temperature is easy to define. We measure the temperature of a glass of water with a household thermometer. We may measure the temperature of a glowing object such as a lamp filament or an iron poker by using an instrument that relates the color of the object and temperature. An iron poker, at room temperature, emits no light, but as it is heated to higher and higher temperatures, it changes in color from dull, barely visible red to a glowing white. We say that the poker is white-hot. Thus we measure and define the temperature of liquids and solids.

But how would you measure the temperature of a gas? At first thought, this seems easy, because we know we can glance at an outdoor thermometer and say that the temperature of the air is 80°, or whatever it happens to be. But what about the temperature of the ionized gas inside a glowing neon tube? The glass walls of the tube are cool to the touch, but inside the tube the neon atoms dash about with astonishing speed, much much faster than the closely packed molecules in a white-hot poker. And what about the temperature of protons in a beam emerging from a cyclotron? Scientists say that an ionized atom moving with a certain speed has an energy of so

many electron volts. But they can also measure this in terms of temperature on a scale in which one electron volt is equivalent to roughly ten thousand degrees Centigrade. On this scale, a 1 Mev (million electron volt) proton has a temperature equivalent of ten billion degrees Centigrade. As we shall see in the next chapter, cyclotrons easily accelerate protons to ten-million electron volts. This corresponds to protons of 100 billion degrees Centigrade, or vastly higher than the temperature of the sun's innermost protons.

A Cornell University physics professor, Dr. Hans Bethe, next tackled the problem of explaining the sun's source of unending energy. In 1938 Bethe was in a much better position to make calculations than Houtermans and Atkinson had been a decade earlier, because experimental scientists had in the meantime come up with so much data about nuclear reactions. Thus Bethe was able to calculate how rapidly protons might combine with one another under conditions existing inside the sun.

Dr. Bethe developed the theory that four protons successively fuse together to form a single atom of helium. This is not accomplished in one fell swoop, but is rather a multiple-stage process in which, first, two hydrogen protons collide and bind themselves together to become an atom of heavy hydrogen, or deuterium; this fused atom of heavy hydrogen is then struck by another proton and helium-3 is formed; finally another proton collision results in the formation of a nucleus of helium-4. The process Bethe envisaged could take place in either of two ways, but both amounted to a synthesis or fusion of four protons, with the release of 27 Mev of energy. The energy that is released comes from the mass "lost" when the four hydrogen atoms fuse into an intimate combination which is lighter than the sum of the individual masses of the H-atoms. The mass "lost" or energy released in a single fusion is small, but because of the enormous amount of hydrogen in the sun, the process occurs frequently enough to keep the sun

blazing hot. Every second about one billion tons of hydrogen undergo fusion! About one million tons of "Einstein mass" are totally converted into energy every second.

Yet this seemingly incredible amount of hydrogen is so small compared with the sun's total supply that the sun will continue to shine at its present rate for billions and billions of years before it runs out of fuel.

If we consider the heat generated per given weight of the sun rather than the total heat produced, we arrive at some rather astonishing facts. On an average, it takes five hundred tons of the sun's mass to produce one hundred watts of heat, the amount given off from a household electric lamp bulb. Even at the sun's center, where the heat is given off at a greater rate, it still takes many tons of the sun's substance to evolve one hundred watts of heat. Actually, the human body—say that of an active teen-ager—generates one hundred times more heat than is generated by an equivalent weight of hydrogen gas in the sun. The explanation is not difficult. In the first place, we are not comparing body temperature with the temperature inside the sun; but rather the rates at which each produces its heat. The sun is almost perfectly insulated by its outer layers of gas, so that even a tiny amount of heat generated at its core, though produced at a much slower rate than in the human body, is kept hot. In other words, the sun's heat is trapped inside its immense mass and leaks out to the surface very gradually. Consequently, the sun continues to build up in temperature; whereas the human body, which is poorly insulated, loses heat rather easily. Even mild exposure to wind suffices to chill a person. One way to look at the problem is to imagine a mass the size of the sun composed of people jammed together as they are in a subway—that is, matter endowed with the heat-producing capacity of an equivalent mass of people. The heat generated would be so great that after a while it would blaze up spectacularly.

The reason heat is evolved so slowly even in the center of the sun is that the hydrogen atoms are at such a low temperature. Roughly twenty million degrees Centigrade may not seem low, but from the standpoint of a nuclear reaction, the equivalent energy of the protons inside the sun's core is only 1,700 electron volts. This is a very low energy for nuclear reactions, since almost all the reactions studied with a cyclotron are measured at energies of millions of volts. Nuclear reactions, especially when we specify thermonuclear reactions, "go" faster at higher energies. This means that deep inside the sun the protons are very weak and fuse together so slowly that it takes millions of years for a hydrogen-helium cycle to occur. That is why our sun doesn't explode like a hydrogen bomb.

Hydrogen bombs release their energy in less than one-millionth of a second. The main reason why such fast reactions can be attained is that heavy and extra-heavy hydrogen are fused in the bomb reaction. Deuterium (double-weight hydrogen) and tritium (triple-weight hydrogen) react violently in contrast to the slow fusion of ordinary or single-weight hydrogen.

In their attempt to make a hydrogen bomb, the experts were up against a cost problem with regard to tritium, and thus it came as a real step ahead when they figured out a way to put a liner of lithium-6 next to the "nuke" in a bomb. The great flash of neutrons released in the explosion of the A-bomb trigger irradiates the lithium liner and gives birth to a burst of tritium atoms. The A-trigger also produces an intense heat wave.

Bomb experts killed two birds with one stone by incorporating the lithium in the form of a chemical compound called lithium deuteride, a compound formed by the synthesis of lithium and heavy hydrogen. They were thus able to bring about the fusion of deuterium and tritium. As we have seen, the fusion process releases energy—in this case, 17.6 Mev for each fusion. This is significantly less than fission energy, but

we must remember that a pound of a light element like lithium contains many more atoms than a pound of a heavy element like uranium and can release more energy.

The energy released in the fusion of hydrogen comes off in the form of high-speed particles, just as in the case of fission. But there is a significant difference, for most of the energy is imparted to the neutron that is produced in the reaction. This neutron dashes off with the lion's share of the fusion energy. It is so speedy that it would tend to flash out into space and not make for a very effective bomb, if the bomb designers had not hit upon an ingenious idea.

They decided to make the runaway neutron do some work in the bomb. They put a heavy jacket of ordinary uranium around the lithium liner. The fast-flying neutrons are trapped in this jacket and there they cause the atoms of U^{238} to fission. The neutrons released in fission, you will recall, will not split U^{238} as readily as they do U^{235}. This is because U^{235} fissions with low-speed neutrons whereas U^{238} does not. Neutrons produced in the chain reaction are not in general sufficiently speedy to fission U^{238}. But, and this is most significant, the neutrons released in hydrogen fusion are fast enough to cause U^{238} to fission.

This means, then, that the superbomb is really a three-stage device. Stage one involves the firing of an atomic bomb trigger. Stage two centers upon the manufacture of tritium from lithium and the fusion of the tritium and heavy hydrogen. Stage three is the fission of ordinary uranium by the fast-fusion neutrons produced in stage two.

All these stages are interrelated by a complex neutron relationship. For example, when U^{238} fissions in stage three, the neutrons produced feed back into the bomb core, causing more fission of the A-trigger and additional production of tritium. In addition, the explosion in stage three creates more heat to produce more fusion. These reactions are so complex and all happen so fast—in one-millionth of a second—that

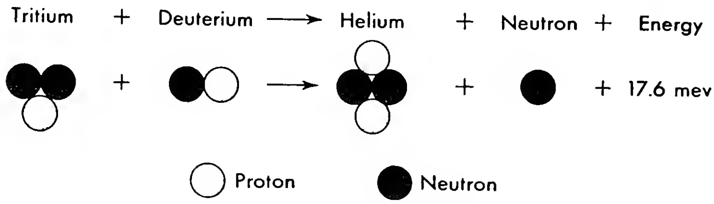
calculation of the bomb's power is exceedingly difficult and must be relegated to whirlwind automatic computers. These electronic brains are capable of lightning-like computation and permit the bomb designers to figure out how a given weapon might perform prior to actual test.

Knowing from the reality of the H-bomb that hydrogen is useful in an explosive thermonuclear reaction, it is natural to ask if hydrogen fusion can be tamed to produce energy useful to man. Is it possible for man to imitate or outdo the sun's energy power?

Before exploring this possibility further, it will help to have clearly in mind why scientists concentrate on hydrogen as a fuel, rather than some other element. Going back to Rutherford's experiments on the scattering of alpha particles, recall that only a very few of the alpha particles penetrated close to the nucleus in the target atom. As the positively charged alpha particles sped toward the positively charged nucleus of the atom, they were strongly repelled by the like electrical forces. The same thing happens when we try to bring together two alpha particles or two hydrogen nuclei or any two nuclei. They resist fusion because of the electrical repulsion of their positively charged cores. The greater the charge on the atomic nucleus, the greater will be the repulsion and hence the difficulty of fusing the two. This means that fusion is easiest for the lighter elements; and hydrogen, with its single proton, is of course the lightest of all.

However, if man attempted to imitate nature's solar process for fusing ordinary hydrogen as fuel, he would be doomed to failure; as we saw earlier, the kind of hydrogen that is present in the sun's interior fuses very slowly, so that a single cubic inch of the central core will evolve only a fraction of a watt of heat energy. The fact of the matter is that ordinary hydrogen is too sluggish a nuclear fuel to support a controlled, man-made fusion reaction. However, as we know, other kinds of hydrogen exist: heavy hydrogen or deuterium, and the

radioactive, extra-heavy form of hydrogen called tritium. Tritium or triple-weight hydrogen can be produced in a nuclear reactor by bombarding lithium with neutrons. Unlike ordinary hydrogen, deuterium and tritium react quickly to



36. Illustrating the fusion of two atoms of hydrogen to form a single atom of helium and a neutron.

create helium; it is this fact that will make controlled fusion power possible. These isotopes are known to undergo the following reactions:

$$_1D^2 + {_1D^2} = {_1H^1} + {_1T^3}$$
$$_1D^2 + {_1D^2} = {_2He^3} + {_0n^1}$$
$$_1D^2 + {_1T^3} = {_2He^4} + {_0n^1}$$
$$_1T^3 + {_1T^3} = {_2He^4} + 2{_0n^1}$$

All these reactions release energy. The first two yield 4.13 and 3.37 Mev respectively, while the last two release 17.58 and 11.32 Mev of energy.

While the energy released by each fusion of hydrogen isotopes is considerably less than the 200 Mev for each fission of a uranium atom, as we noted earlier in the case of lithium, the number of atoms in a pound of hydrogen is very much greater than the number of atoms in a pound of uranium. A pound of deuterium, for instance, releases roughly three times as much energy as a pound of uranium. Converted into the energy content of the heavy hydrogen in a cup of water, this amounts to the heat equivalent of fifty pounds of coal. The supply of heavy hydrogen is practically without limit since the lakes and oceans on our planet contain inexhaustible re-

serves of water. Thus, if man can extract hydrogen fusion energy, he has at hand an unlimited new supply of fuel.

The goal of hydrogen power is tempting for more than just this reason. Hydrogen fusion produces no residual radioactive fragments, so the radiation hazard of uranium fission products is not present in this new type of power source. Furthermore, because of the nature of the reactor that will probably be used to produce fusion power, there is no danger of a runaway explosion, such as can occur in certain types of uranium power plants. In addition, there is the enticing prospect that it may be possible to derive energy from a fusion reactor directly, in the form of electrical power.

Attractive as these prospects appear, one has to consider the huge difficulties that stand in the road toward attaining fusion power. The basic fuel, deuterium, is no problem, since heavy water can be produced in hundred-ton lots and is readily available commercially at $28 per pound. And there is no problem in obtaining pure deuterium gas from the heavy water. The fundamental problem is so to design a reactor that ionized deuterium, or hydrogen plasma as it is called, can be brought to sufficiently high speed for fusion to take place. This requires that a temperature above one hundred million degrees Centigrade be attained.

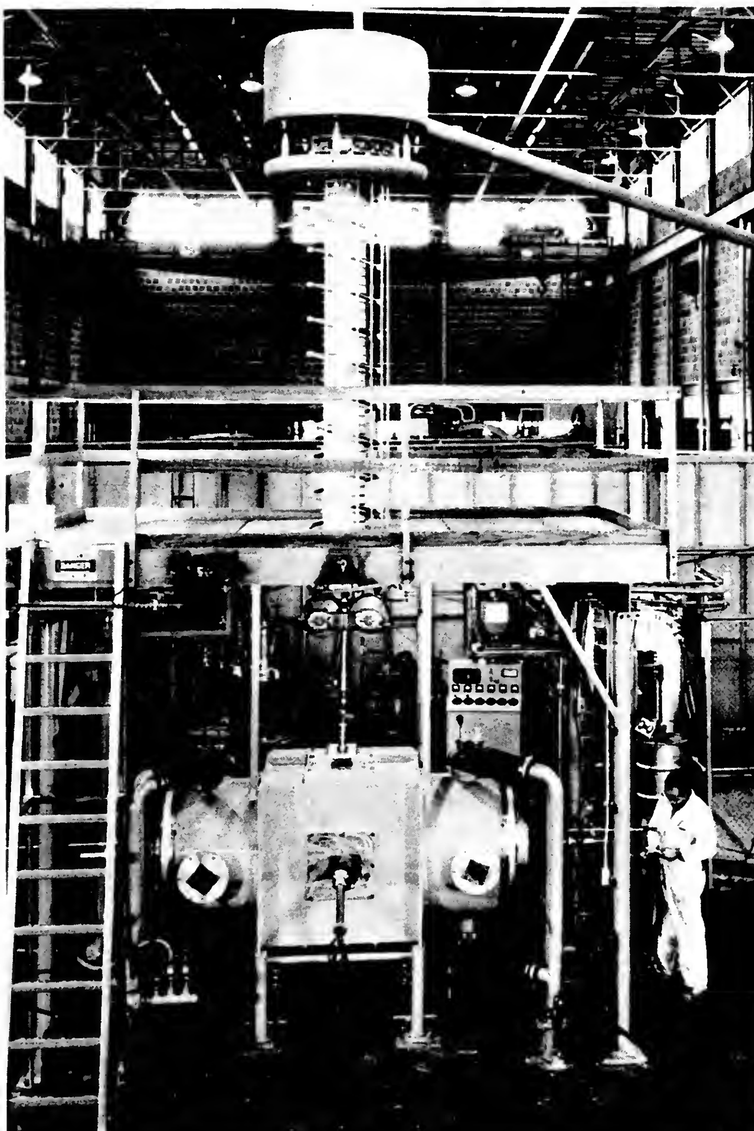
Scientists in many countries are hard at work designing machines that will use electric and magnetic fields to squeeze hydrogen plasma together or "pinch" it. The basic idea was set forth in 1934 by the American physicist, W. H. Bennett. He suggested that charged particles of hydrogen moving in a stream would constitute an electrical current that should induce its own magnetic fields; this, in turn, would act to pinch the plasma together, bringing the individual ions into collision with each other. The more violent the collisions (i.e., the "hotter" the pinch) and the more frequent they are, the greater is the probability that fusion will occur.

Unfortunately, the phenomenon just described is not very

easy to control or stabilize. In the United States, the Atomic Energy Commission established Project Sherwood for the purpose of bringing about the controlled release of fusion power. The research work, begun on a modest budget in 1951, expanded to a vigorous program in 1959, backed by a forty-million-dollar annual budget. A variety of experimental devices for studying the "pinch" effect have been built at the Los Alamos Scientific Laboratory, of which the Perhapstron is an example. Hydrogen ions are circulated in a doughnut-shaped vacuum tube and constricted by an electrical current into a narrow column inside this chamber.

A "Magnetic Mirror" device represents a different approach to the fusion problem adopted by scientists at the University of California's Livermore Laboratory. Instead of a doughnut chamber, a straight tube is employed and the hydrogen plasma is "trapped" by intense magnetic fields and "reflected" back from one end of the tube into the center of the chamber. Still another line of approach is shown in the illustration. Here at Oak Ridge, scientists are studying fusion possibilities by hurling heavy hydrogen molecules downward into a reaction chamber where they are ionized by an electric arc and then subjected to intense magnetic forces. A more ambitious and larger-scale approach to fusion power is under way at Princeton University, where a Stellerator is being constructed. Magnetic forces from a thick magnetic coil that is wrapped around a figure-8-shaped vacuum chamber center the hydrogen ions in the chamber. This unusual container is designed to keep the hydrogen ions from straying out to the wall and giving up their energy. Fusion power can be attained only if the plasma can be kept isolated from contact with the container.

Obviously, no structural container can hold anything so hot as this fiery plasma. Instead, scientists propose to contain the plasma by means of magnetic fields which force the ions to stay in a restricted space, i.e., a kind of "magnetic bottle."



37. The Oak Ridge Fusion research device designed to probe hydrogen fusion on a laboratory scale. (Oak Ridge National Laboratory)

However, there is the serious problem of designing such a magnetic "container" so that it is substantially leakproof. Any small leak would allow the hot plasma to squirt out to the tube wall and cool off, thus ruining chances of attaining the high temperatures necessary for fusion. Experiments in the United States have produced plasma at a temperature of about ten million degrees Centigrade.

Fusion research is also going on in Russia, Britain, Sweden, Germany, Japan and many other countries. The British have pioneered in this new field of research and have constructed rather large machines. All machines concentrate on using deuterium as the reacting substance, although later experiments may be done with tritium. However, tritium is more difficult to handle experimentally because of the radiation hazard and the contamination of the equipment.

If one selects pure deuterium as the nuclear fuel for fusion power, there is the attractive prospect that, since two-thirds of the energy comes off in the form of charged particles, it might be possible to convert this directly into electrical energy. Picturing the way a piston functions in a steam engine, one may think of moving plasma working against magnetic fields, and electrical circuits drawing off the energy. With a mixture of deuterium and tritium, the majority of the energy is carried off by the neutrons. A blanket of liquid lithium might be used to absorb the neutrons and convert their energy into heat and at the same time generate useful tritium as the lithium atoms are fissioned. Thus fusion power would be used to produce heat external to the plasma and this heat would then be used for the purposes of producing more power.

The possibility of fusion power is raised at a time when uranium power plants are being engineered to produce power on a basis competitive with conventional fuels. Rising coal costs in England have provided the British with a strong incentive to replace coal with uranium and they have devoted tremendous effort to building uranium power stations. Now

there is the question whether uranium power is not obsolete before it is even fully developed. Will not fusion of hydrogen replace uranium fission as man's source of energy? Ultimately, it seems clear that hydrogen fusion will be developed to the point where it is attractive for some applications, but this new source of power is in its technological infancy and it is too early to predict when it will assume its place in the sun. However, it can be said that many scientists who are working on this ultimate fuel are optimistic that they will be able to solve the very formidable problems that lie ahead. Furthermore, they feel that in their explorations of high-temperature plasmas and intense magnetic fields they will learn many new facts about atoms and the cosmos. Indeed, some scientists believe that even if hydrogen power should never succeed, should man be frustrated in his attempt to outdo the sun, he will gather rich dividends in fundamental knowledge, and the research will have been worth while. But the hope is that the quest for fusion power will bring to mankind an unlimited source of power to heat homes, light cities and power factories for millions of years to come.

Mrs. Enrico Fermi gives in colorful detail her personal account of the first nuclear chain reaction at the University of Chicago squash courts.

14 Success

Laura Fermi

Chapter from her book, *Atoms in the Family,* published in 1954.

Meanwhile Herbert Anderson and his group at the Met. Lab. had also been building small piles and gathering information for a larger pile from their behavior. The best place Compton had been able to find for work on the pile was a squash court under the West Stands of Stagg Field, the University of Chicago stadium. President Hutchins had banned football from the Chicago campus, and Stagg Field was used for odd purposes. To the west, on Ellis Avenue, the stadium is closed by a tall gray-stone structure in the guise of a medieval castle. Through a heavy portal is the entrance to the space beneath the West Stands. The Squash Court was part of this space. It was 30 feet wide, twice as long, and over 26 feet high.

The physicists would have liked more space, but places better suited for the pile, which Professor Compton had hoped he could have, had been requisitioned by the expanding armed forces stationed in Chicago. The physicists were to be contented with the Squash Court, and there Herbert Anderson had started assembling piles. They were still "small piles," because material flowed to the West Stands at a very slow, if steady, pace. As each new shipment of crates arrived, Herbert's spirits rose. He loved working and was of impatient temperament. His slender, almost delicate, body had unsuspected resilience and endurance. He could work at all hours and drive his associates to work along with his same intensity and enthusiasm.

A shipment of crates arrived at the West Stands on a Saturday afternoon, when the hired men who would normally unpack them were not working. A university professor, older by several years than Herbert, gave a look at the crates and said lightly: "Those fellows will unpack them Monday morning."

"Those fellows, Hell! We'll do them now," flared up Herbert, who

had never felt inhibited in the presence of older men, higher up in the academic hierarchy. The professor took off his coat, and the two of them started wrenching at the crates.

Profanity was freely used at the Met. Lab. It relieved the tension built up by having to work against time. Would Germany get atomic weapons before the United States developed them? Would these weapons come in time to help win the war? These unanswered questions constantly present in the minds of the leaders in the project pressed them to work faster and faster, to be tense, and to swear.

Success was assured by the spring. A small pile assembled in the Squash Court showed that all conditions—purity of materials, distribution of uranium in the graphite lattice—were such that a pile of critical size would chain-react.

"It could be May, or early June at latest," Enrico told me, as we recently reminisced about the times of the Met. Lab. "I remember I talked about that experiment on the Indiana dunes, and it was the first time I saw the dunes. You were still in Leonia. I went with a group from the Met. Lab. I liked the dunes: it was a clear day, with no fog to dim colors. . . ."

"I don't want to hear about the dunes," I said. "Tell me about that experiment."

"I like to swim in the lake, . . ." Enrico paid no attention to my remark. I knew that he enjoyed a good swim, and I could well imagine him challenging a group of younger people, swimming farther and for a longer time than any of them, then emerging on the shore with a triumphant grin.

"Tell me about that experiment," I insisted.

"We came out of the water, and we walked along the beach."

I began to feel impatient. He did not have to mention the walk. He always walks after swimming, dripping wet, water streaming from his hair. In 1942 there was certainly much more hair on his head to shed water, not just the little fringe on the sides and on the back that there is now, and it was much darker.

". . . and I talked about the experiment with Professor Stearns. The two of us walked ahead of the others on the beach. I remember our efforts to speak in such a way that the others would not understand. . . ."

"Why? Didn't everyone at the Met. Lab. know that you were building piles?"

"They knew we built piles. They did not know that at last we had the certainty that a pile would work. The fact that a chain reaction was feasible remained classified material for a while. I could talk freely with Stearns because he was one of the leaders."

"If you were sure a larger pile would work, why didn't you start it at once?"

"We did not have enough materials, neither uranium nor graphite. Procurement of uranium metal was always an obstacle. It hampered progress."

While waiting for more materials, Herbert Anderson went to the Goodyear Tire and Rubber Company to place an order for a square balloon. The Goodyear people had never heard of square balloons, they did not think they could fly. At first they threw suspicious glances at Herbert. The young man, however, seemed to be in full possession of his wits. He talked earnestly, had figured out precise specifications, and knew exactly what he wanted. The Goodyear people promised to make a square balloon of rubberized cloth. They delivered it a couple of months later to the Squash Court. It came neatly folded, but, once unfolded, it was a huge thing that reached from floor to ceiling.

The Squash Court ceiling could not be pushed up as the physicists would have liked. They had calculated that their final pile ought to chain-react somewhat before it reached the ceiling. But not much margin was left, and calculations are never to be trusted entirely. Some impurities might go unnoticed, some unforeseen factor might upset theory. The critical size of the pile might not be reached at the ceiling. Since the physicists were compelled to stay within that very concrete limit, they thought of improving the performance of the pile by means other than size.

The experiment at Columbia with a canned pile had indicated that such an aim might be attained by removing the air from the pores of the graphite. To can as large a pile as they were to build now would be impracticable, but they could assemble it inside a square balloon and pump the air from it if necessary.

The Squash Court was not large. When the scientists opened the balloon and tried to haul it into place, they could not see its top

from the floor. There was a movable elevator in the room, some sort of scaffolding on wheels that could raise a platform. Fermi climbed onto it, let himself be hoisted to a height that gave him a good view of the entire balloon, and from there he gave orders:

"All hands stand by!"

"Now haul the rope and heave her!"

"More to the right!"

"Brace the tackles to the left!"

To the people below he seemed an admiral on his bridge, and "Admiral" they called him for a while.

When the balloon was secured on five sides, with the flap that formed the sixth left down, the group began to assemble the pile inside it. Not all the material had arrived, but they trusted that it would come in time.

From the numerous experiments they had performed so far, they had an idea of what the pile should be, but they had not worked out the details, there were no drawings nor blueprints and no time to spare to make them. They planned their pile even as they built it. They were to give it the shape of a sphere of about 26 feet in diameter, supported by a square frame, hence the square balloon.

The pile supports consisted of blocks of wood. As a block was put in place inside the balloon, the size and shape of the next were figured. Between the Squash Court and the near-by carpenter's shop there was a steady flow of boys, who fetched finished blocks and brought specifications for more on bits of paper.

When the physicists started handling graphite bricks, everything became black. The walls of the Squash Court were black to start with. Now a huge black wall of graphite was going up fast. Graphite powder covered the floor and made it black and as slippery as a dance floor. Black figures skidded on it, figures in overalls and goggles under a layer of graphite dust. There was one woman among them, Leona Woods; she could not be distinguished from the men, and she got her share of cussing from the bosses.

The carpenters and the machinists who executed orders with no knowledge of their purpose and the high-school boys who helped lay bricks for the pile must have wondered at the black scene. Had they been aware that the ultimate result would be an atomic bomb, they might have renamed the court Pluto's Workshop or Hell's Kitchen.

To solve difficulties as one meets them is much faster than to try to foresee them all in detail. As the pile grew, measurements were taken and further construction adapted to results.

The pile never reached the ceiling. It was planned as a sphere 26 feet in diameter, but the last layers were never put into place. The sphere remained flattened at the top. To make a vacuum proved unnecessary, and the balloon was never sealed. The critical size of the pile was attained sooner than was anticipated.

Only six weeks had passed from the laying of the first graphite brick, and it was the morning of December 2.

Herbert Anderson was sleepy and grouchy. He had been up until two in the morning to give the pile its finishing touches. Had he pulled a control rod during the night, he could have operated the pile and have been the first man to achieve a chain reaction, at least in a material, mechanical sense. He had a moral duty not to pull that rod, despite the strong temptation. It would not be fair to Fermi. Fermi was the leader. He had directed research and worked out theories. His were the basic ideas. His were the privilege and the responsibility of conducting the final experiment and controlling the chain reaction.

"So the show was all Enrico's, and he had gone to bed early the night before," Herbert told me years later, and a bit of regret still lingered in his voice.

Walter Zinn also could have produced a chain reaction during the night. He, too, had been up and at work. But he did not care whether he operated the pile or not; he did not care in the least. It was not his job.

His task had been to smooth out difficulties during the pile construction. He had been some sort of general contractor: he had placed orders for material and made sure that they were delivered in time; he had supervised the machine shops where graphite was milled; he had spurred others to work faster, longer, more efficiently. He had become angry, had shouted, and had reached his goal. In six weeks the pile was assembled, and now he viewed it with relaxed nerves and with that vague feeling of emptiness, of slight disorientation, which never fails to follow completion of a purposeful task.

There is no record of what were the feelings of the three young men who crouched on top of the pile, under the ceiling of the square

balloon. They were called the "suicide squad." It was a joke, but perhaps they were asking themselves whether the joke held some truth. They were like firemen alerted to the possibility of a fire, ready to extinguish it. If something unexpected were to happen, if the pile should get out of control, they would "extinguish" it by flooding it with a cadmium solution. Cadmium absorbs neutrons and prevents a chain reaction.

Leona Woods, the one girl in that group of men, was calm and composed; only the intensity of her deep dark eyes revealed the extent of her alertness.

Among the persons who gathered in the Squash Court on that morning, one was not connected with the Met. Lab.—Mr. Crawford H. Greenewalt of E. I. duPont de Nemours, who later became the president of the company. Arthur Compton had led him there out of a near-by room where, on that day, he and other men from his company happened to be holding talks with top Army officers.

Mr. Greenewalt and the duPont people were in a difficult position, and they did not know how to reach a decision. The Army had taken over the Uranium Project on the previous August and renamed it Manhattan District. In September General Leslie R. Groves was placed in charge of it. General Groves must have been of a trusting nature: before a chain reaction was achieved, he was already urging the duPont de Nemours Company to build and operate piles on a production scale.

In a pile, Mr. Greenewalt was told, a new element, plutonium, is created during uranium fission. Plutonium would probably be suited for making atomic bombs. So Greenewalt and his group had been taken to Berkeley to see the work done on plutonium, and then flown to Chicago for more negotiations with the Army.

Mr. Greenewalt was hesitant. Of course his company would like to help win the war! But piles and plutonium!

With the Army's insistent voice in his ears, Compton, who had attended the conference, decided to break the rules and take Mr. Greenewalt to witness the first operation of a pile.

They all climbed onto the balcony at the north end of the Squash Court; all, except the three boys perched on top of the pile and except a young physicist, George Weil, who stood alone on the floor

by a cadmium rod that he was to pull out of the pile when so instructed.

And so the show began.

There was utter silence in the audience, and only Fermi spoke. His gray eyes betrayed his intense thinking, and his hands moved along with his thoughts.

"The pile is not performing now because inside it there are rods of cadmium which absorb neutrons. One single rod is sufficient to prevent a chain reaction. So our first step will be to pull out of the pile all control rods, but the one that George Weil will man." As he spoke others acted. Each chore had been assigned in advance and rehearsed. So Fermi went on speaking, and his hands pointed out the things he mentioned.

"This rod, that we have pulled out with the others, is automatically controlled. Should the intensity of the reaction become greater than a pre-set limit, this rod would go back inside the pile by itself.

"This pen will trace a line indicating the intensity of the radiation. When the pile chain-reacts, the pen will trace a line that will go up and up and that will not tend to level off. In other words, it will be an exponential line.

"Presently we shall begin our experiment. George will pull out his rod a little at a time. We shall take measurements and verify that the pile will keep on acting as we have calculated.

"Weil will first set the rod at thirteen feet. This means that thirteen feet of the rod will still be inside the pile. The counters will click faster and the pen will move up to this point, and then its trace will level off. Go ahead, George!"

Eyes turned to the graph pen. Breathing was suspended. Fermi grinned with confidence. The counters stepped up their clicking; the pen went up and then stopped where Fermi had said it would. Greenewalt gasped audibly. Fermi continued to grin.

He gave more orders. Each time Weil pulled the rod out some more, the counters increased the rate of their clicking, the pen raised to the point that Fermi predicted, then it leveled off.

The morning went by. Fermi was conscious that a new experiment of this kind, carried out in the heart of a big city, might become a potential hazard unless all precautions were taken to make sure that at all times the operation of the pile conformed closely with the

results of the calculations. In his mind he was sure that if George Weil's rod had been pulled out all at once, the pile would have started reacting at a leisurely rate and could have been stopped at will by reinserting one of the rods. He chose, however, to take his time and be certain that no unforeseen phenomenon would disturb the experiment.

It is impossible to say how great a danger this unforeseen element constituted or what consequences it might have brought about. According to the theory, an explosion was out of the question. The release of lethal amounts of radiation through an uncontrolled reaction was improbable. Yet the men in the Squash Court were working with the unknown. They could not claim to know the answers to all the questions that were in their minds. Caution was welcome. Caution was essential. It would have been reckless to dispense with caution.

So it was lunch time, and, although nobody else had given signs of being hungry, Fermi, who is a man of habits, pronounced the now historical sentence:

"Let's go to lunch."

After lunch they all resumed their places, and now Mr. Greenewalt was decidedly excited, almost impatient.

But again the experiment proceeded by small steps, until it was 3:20.

Once more Fermi said to Weil:

"Pull it out another foot"; but this time he added, turning to the anxious group in the balcony: "This will do it. Now the pile will chain-react."

The counters stepped up; the pen started its upward rise. It showed no tendency to level off. A chain reaction was taking place in the pile.

Leona Woods walked up to Fermi and in a voice in which there was no fear she whispered: "When do we become scared?"

Under the ceiling of the balloon the suicide squad was alert, ready with their liquid cadmium: this was the moment. But nothing much happened. The group watched the recording instruments for 28 minutes. The pile behaved as it should, as they all had hoped it would, as they had feared it would not.

The rest of the story is well known. Eugene Wigner, the Hun-

garian-born physicist who in 1939 with Szilard and Einstein had alerted President Roosevelt to the importance of uranium fission, presented Fermi with a bottle of Chianti. According to an improbable legend, Wigner had concealed the bottle behind his back during the entire experiment.

(See letter on pp. 132–133)

All those present drank. From paper cups, in silence, with no toast. Then all signed the straw cover on the bottle of Chianti. It is the only record of the persons in the Squash Court on that day.

The group broke up. Some stayed to round up their measurements and put in order the data gathered from their instruments. Others went to duties elsewhere. Mr. Greenewalt hastened to the room where his colleagues were still in conference with the military. He announced, all in one breath, that Yes, it would be quite all right for their company to go along with the Army's request and start to build piles. Piles were wonderful objects that performed with the precision of a Swiss watch, and, provided that the advice of such competent scientists as Fermi and his group were available, the duPont company was certainly taking no undue risk.

Arthur Compton placed a long-distance call to Mr. Conant of the Office of Scientific Research and Development at Harvard.

"The Italian Navigator has reached the New World," said Compton as soon as he got Conant on the line.

"And how did he find the natives?"

"Very friendly."

Here the official story ends, but there is a sequel to it, which started on that same afternoon when a young physicist, Al Wattemberg, picked up the empty Chianti bottle from which all had drunk. With the signatures on its cover, it would make a nice souvenir. In subsequent years Al Wattemberg did his share of traveling, like any other physicist, and the bottle followed him. When big celebrations for the pile's tenth anniversary were planned at the University of Chicago, the bottle and Al Wattemberg were both in Cambridge, Massachusetts. Both, Al promised, would be in Chicago on December 2.

It so happened, however, that a little Wattemberg decided to come into this world at about that time, and Al could not attend the celebrations. So he shipped his bottle, and, because he wanted to make doubly sure that it would not be broken, he insured it for a

(continued on p. 134)

Albert Einstein
Old Grove Rd.
Nassau Point
Peconic, Long Island

August 2nd, 1939

F.D. Roosevelt,
President of the United States,
White House
Washington, D.C.

Sir:

Some recent work by E.Fermi and L. Szilard, which has been communicated to me in manuscript, leads me to expect that the element uranium may be turned into a new and important source of energy in the immediate future. Certain aspects of the situation which has arisen seem to call for watchfulness and, if necessary, quick action on the part of the Administration. I believe therefore that it is my duty to bring to your attention the following facts and recommendations:

In the course of the last four months it has been made probable - through the work of Joliot in France as well as Fermi and Szilard in America - that it may become possible to set up a nuclear chain reaction in a large mass of uranium,by which vast amounts of power and large quantities of new radium-like elements would be generated. Now it appears almost certain that this could be achieved in the immediate future.

This new phenomenon would also lead to the construction of bombs, and it is conceivable - though much less certain - that extremely powerful bombs of a new type may thus be constructed. A single bomb of this type, carried by boat and exploded in a port, might very well destroy the whole port together with some of the surrounding territory. However, such bombs might very well prove to be too heavy for transportation by air.

The United States has only very poor ores of uranium in moderate quantities. There is some good ore in Canada and the former Czechoslovakia, while the most important source of uranium is Belgian Congo.

In view of this situation you may think it desirable to have some permanent contact maintained between the Administration and the group of physicists working on chain reactions in America. One possible way of achieving this might be for you to entrust with this task a person who has your confidence and who could perhaps serve in an inofficial capacity. His task might comprise the following:

a) to approach Government Departments, keep them informed of the further development, and put forward recommendations for Government action, giving particular attention to the problem of securing a supply of uranium ore for the United States;

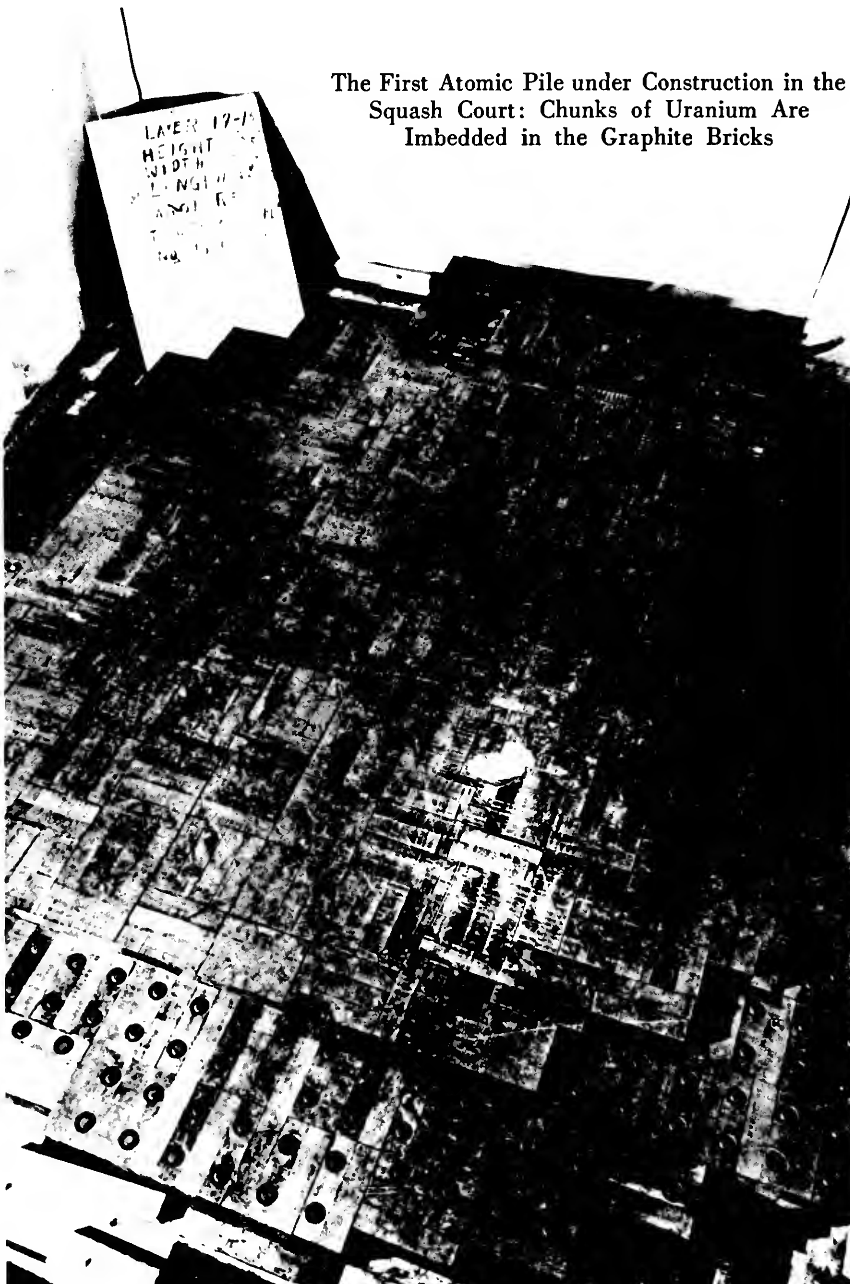
b) to speed up the experimental work, which is at present being carried on within the limits of the budgets of University laboratories, by providing funds, if such funds be required, through his contacts with private persons who are willing to make contributions for this cause, and perhaps also by obtaining the co-operation of industrial laboratories which have the necessary equipment.

I understand that Germany has actually stopped the sale of uranium from the Czechoslovakian mines which she has taken over. That she should have taken such early action might perhaps be understood on the ground that the son of the German Under-Secretary of State, von Weizsäcker, is attached to the Kaiser-Wilhelm-Institut in Berlin where some of the American work on uranium is now being repeated.

Yours very truly,
A. Einstein
(Albert Einstein)

thousand dollars. It is not often that an empty bottle is considered worth so much money, and newspaper men on the lookout for sensation gave the story a prominent position in the press.

A couple of months later the Fermis and a few other physicists received a present: a case of Chianti wine. An importer had wished to acknowledge his gratitude for the free advertisement that Chianti had received.



The First Atomic Pile under Construction in the Squash Court: Chunks of Uranium Are Imbedded in the Graphite Bricks

Until now, power from nuclear reactors has been too expensive for widespread civilian use in this country. But today electricity from such reactors is economically competitive and is projected to become much cheaper.

---

15 The Nuclear Energy Revolution

Alvin M. Weinberg and Gale Young

Excerpt from a lecture given at the National Academy of Sciences in 1966.

Twenty-four years have passed since Fermi and his co-workers at Chicago achieved the first nuclear chain reaction. During most of these years nuclear power for civilian use has been too expensive and experimental in nature to play much of a role in our economy, but during the past couple of years the situation has changed. Nuclear reactors now appear to be the cheapest of all sources of energy. We believe, and this belief is shared by many others working in nuclear energy, that we are only at the beginning, and that nuclear energy will become cheap enough to influence drastically the many industrial processes that use energy. If nuclear energy does not, as H. G. Wells put it in 1914, create "A World Set Free," it will nevertheless affect much of the economy of the coming generation. It is this Nuclear Energy Revolution, based upon the permanent and ubiquitous availability of cheap nuclear power, about which we shall speculate.

Our outlook is admittedly optimistic; yet optimism in nuclear energy seems justified. In 1955, at the first International Conference for the Peaceful Uses of Atomic Energy, in Geneva, some American authorities were chided for predicting nuclear power priced at 4–5 mills per kilowatt hour (kwh). Today TVA has announced that it expects to generate power from its 2200-megawatt (Mw) Browns Ferry boiling-water nuclear plant at 2.4 mills/kwh. Even if the Browns Ferry plant were operated by a private utility, the electricity at the bus bar would cost less than 3.5 mills/kwh. We are very hopeful that still lower costs will be achieved in the future with breeder reactors.

Cheap Nuclear Energy Is Close at Hand.—The economic breakthrough in nuclear energy came in 1963 when the Jersey Central Power and Light Company contracted with the General Electric Company to construct the Oyster Creek boiling-water nuclear power plant. At its expected electrical output of 620-Mw the capital cost of this plant is $110/kw or the same as that for a coal-fired power plant of the same size at the same location.[1] The announcement of Oyster Creek was at first regarded by many as a sort of fluke. But Oyster Creek was followed by a succession of orders for large light-water-cooled power plants, so that now there are 29 com-

TABLE 1
RECENT SALES OF WATER REACTORS

Plant	Utility	Nominal Mw	Manufacturer
Oyster Creek	Jersey Central	515	General Electric
San Onofre	Southern California Edison	429	Westinghouse
Nine Mile Point	Niagara Mohawk	500	General Electric
Haddam Neck	Connecticut Yankee	463	Westinghouse
Dresden 2	Commonwealth Edison	755	General Electric
—	Boston Edison	600	General Electric
Millstone Point	Northeast Utilities	549	General Electric
Brookwood	Rochester Gas & Electric	420	Westinghouse
Indian Point 2	Consolidated Edison	873	Westinghouse
Turkey Point 3	Florida Power & Light	652	Westinghouse
Turkey Point 4	Florida Power & Light	652	Westinghouse
Dresden 3	Commonwealth Edison	810	General Electric
Robinson	Carolina Power & Light	760	Westinghouse
Palisades	Consumers Power Company	810	Combustion Engr.
Point Beach	Wisconsin Michigan Power	480	Westinghouse
Quad Cities 1 and 2	Commonwealth Edison and Iowa-Illinois G & E	2 × 810	General Electric
Monticello	Northern States Power Co.	540	General Electric
Browns Ferry	TVA	2 × 1100	General Electric
Vernon	Vermont Yankee	540	General Electric
Keowee Dam	Duke Power Company	2 × 820	Babcock and Wilcox
Peach Bottom 2	Philadelphia Electric	2 × 1100	General Electric
Delaware Valley	Public Service Electric & Gas of New Jersey	1000	Westinghouse
Surry	Virginia Electric Power Co.	2 × 800	Westinghouse
Boston	Boston Edison	600	General Electric

mitments for construction of large nuclear power reactors in the United States (Table 1). More than half of the large station generating capacity ordered in recent months is scheduled to be nuclear.

None of the plants listed in Table 1 are as yet operating. Oyster Creek will go on the line early in 1968. The optimism expressed in the many purchases of light-water-moderated and cooled reactors is based partly upon our generally good experience with such reactors in the nuclear navy, and partly upon the operating experience with such power plants as the Yankee pressurized-water reactor (175 Mw) and the Dresden 1 boiling-water reactor (200 Mw). Yankee, for example, has been generating electricity for five years, and during the past year has been available for generation 76 per cent of the time. Dresden 1 has operated for six years, and during the past year has been available 83 per cent of the time.

In some ways it is surprising that the world's cheapest nuclear reactors should derive from the original pressurized-water line used to power the *Nautilus*. Pressurized water was chosen for the *Nautilus* not because it seemed to be a path to cheap nuclear energy, but rather because such reactors, being moderated by hydrogen and fueled with enriched uranium, are relatively compact. If anything, the early reactor designers viewed these systems as being rather expensive. And in countries other than the United States and the Soviet Union, the main-line reactors utilize natural uranium and either graphite or heavy water as moderator.

But the early designers failed to appreciate the extent to which the extraordinary success of the gaseous diffusion plants would reduce the price of U^{235}. In 1948, when the *Nautilus* was designed, U^{235} cost about \$35/gm. Today it costs \$12/gm, which is only four times its price as unseparated isotope in ore costing \$8/lb of U_3O_8! This remarkable reduction in the cost of separating U^{235}, more than any other single factor, underlies the economic success of the American water-moderated reactors.

The fuel cycle in a reactor like Browns Ferry that burns enriched uranium costs only 1.25 mills/kwh, which is appreciably lower than coal even in cheap coal country (Table 2).

The American reactors, being compact, were expected to be cheaper to build than the large graphite or heavy-water reactors that use natural uranium. But prior to Oyster Creek it was not clear how cheap a reactor could be, especially if its output were large enough. It was R. P. Hammond who first stressed the principle that a nuclear reactor ought to scale rather favorably. Thus, although the total cost of a large nuclear reactor will be greater than that of a smaller one, the cost per kilowatt of the large reactor should be less than that of the smaller one. Hammond's contention has been amply confirmed by the price estimates published, for example, by the General Electric Company. Figure 1 shows that the cost per kilowatt of a 200-Mw boiling-water reactor (BWR) is around $180/kw, whereas the cost per kilowatt of a 1000-Mw BWR is only $110/kw. All the new, competitive nuclear power plants are large, and they capitalize on the advantage of size.

The Necessity for Breeders.—Nuclear power at 2.4 mills/kwh at Browns Ferry is a remarkable achievement, but it is not remarkable enough to serve as the basis for a Nuclear Energy Revolution. In the first place, we are hopeful that breeder reactors can shave another mill off the cost and thus perhaps provide the basis for new heavy chemical and other industries. In the second place, the light-water reactors burn only a small fraction of all the natural uranium mined to fuel them; thus such reactors will rapidly use all the U. S. low-priced reserves of uranium ore, and the price of nuclear energy will rise as we are obliged to burn more expensive ores. This is illustrated in Figure 2, based by Dietrich[3] on estimates made a few years ago by the Atomic Energy Commission of U. S. ore reserves and reactors to be built.[4] Since then, ore prospecting has been resumed, but water reactor sales are outrunning the estimates.

We therefore find ourselves in a serious dilemma. The current great success of nuclear energy is making our economy increasingly dependent upon nuclear power. But as we turn to nuclear energy we shall exhaust our low-grade ore reserves. By the time (say in 1990) we have become very heavily committed to nuclear energy, its price will probably begin to rise significantly.

Of course we shall find more low-cost ore. But eventually even this will be insufficient, especially if our power requirements continue to grow. If we are to forestall a major economic power crisis, say 20 years from now, we shall have to learn how to utilize not 1 per cent or so of the raw materials (uranium and thorium) for fuel, but much more—hopefully close to 100 per cent. Should we learn how to burn a large fraction of the uranium and thorium, we would gain in three respects: we would forestall a serious rise in the cost of power; we would reduce the fuel cycle cost of a reactor, since in effect we would be burning the abundant and very cheap U^{238} or Th^{232}, not the costly U^{235}; and we would make available, at relatively small economic penalty, the vast residual amounts of uranium and thorium in the earth's crust. To anticipate our conclusion, we could hope to achieve power costs of only 1.5 mills/kwh in publicly owned stations, and we could foresee maintaining this low cost essentially forever. It is this prospect, and what it implies for energy-consuming industrial processes, that warrants our using the extravagant phrase "The Nuclear Energy Revolution."

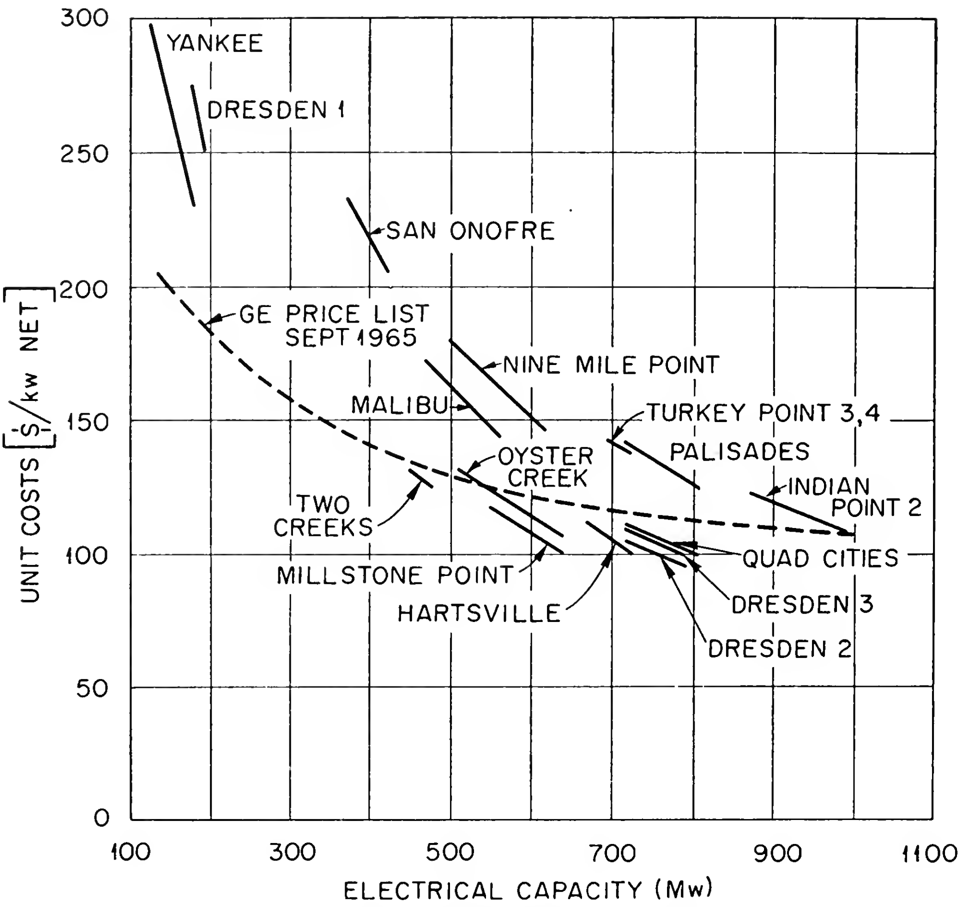


FIG. 1.—Cost of nuclear electric plants. The length of each short-line segment represents the uncertainty in the ultimate output of each reactor. The values shown are mostly manufacturers' "turn-key" prices, and do not in many cases include all the customers' costs. Complete data are usually not available.

TABLE 2

SOME CURRENT POWER COST ESTIMATES[1, 2]

	Oyster Creek nuclear	TVA nuclear	TVA coal
Investment ($/kw)	116*	116†	117
Capacity assumed	Expected stretch, 620 Mw	Guaranteed 1100 Mw	
Plant life (yr)	30	35	35
Fixed charge rate (%/yr)	10	5.7	5.7
Load factor (%)	88	85	85
Period covered (yr)	First 10	First 12	First 12
Capital charges (mills/kwh)	1.5	0.89	0.90
Operation, maintenance, insurance (mills/kwh)	0.48	0.23	0.24
Fuel cycle (mills/kwh)	1.67	1.25	1.69
Total power cost (mills/kwh)	3.65	2.37	2.83

* Includes $4/kw transmission and $2/kw working capital other than fuel.
† Reduces by $9/kw if anticipated stretch is realized.

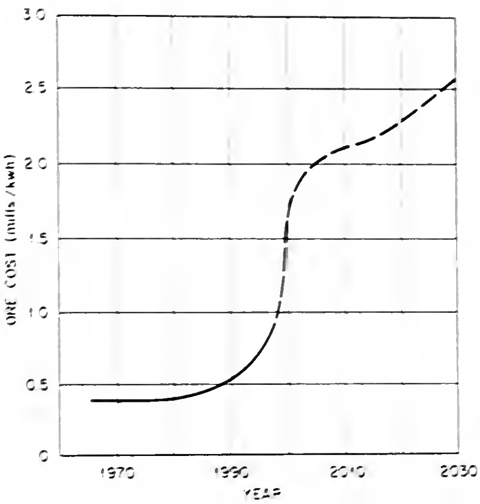


FIG. 2.—Ore costs for H_2O reactors with plutonium recycle.

[1] Jersey Central Power and Light Company, "Report on economic analysis for Oyster Creek nuclear electric generating station," *Nuclear News*, 7, no. 4, Special Supplement (April 1964). The station being built is a little less expensive than the one analyzed in the report.

[2] Tennessee Valley Authority, *Comparison of Coal-Fired and Nuclear Power Plants for the TVA System* (Chattanooga, Tenn.: Office of Power, June 1966).

[3] Dietrich, J. R., "Efficient utilization of nuclear fuels," *Power Reactor Technology*, 6, no. 4, 34 (Fall 1963), U.S. Atomic Energy Commission Division of Technical Information, Oak Ridge, Tennessee.

[4] U. S. Atomic Energy Commission, *Civilian Nuclear Power: A Report to the President—1962, and Appendices* (Oak Ridge, Tenn.: U.S. Atomic Energy Commission Division of Technical Information Extension, 1962).

In the study of elementary particles, new conservation laws have been discovered that are indispensible for making prediction or building theory.

16 Conservation Laws

Kenneth W. Ford

Chapter from his book, *The World of Elementary Particles*, published in 1963.

In a slow and subtle, yet inexorable, way conservation laws have moved in the past few centuries from the role of interesting sidelight in physics to the most central position. What little we now understand about the interactions and transformations of particles comes in large part through certain conservation laws which govern elementary-particle behavior.

A conservation law is a statement of constancy in nature. If there is a room full of people, say, at a cocktail party, and no one comes in or leaves, we can say that there is a law of conservation of the number of people; that number is a constant. This would be a rather uninteresting law. But suppose the conservation law remained valid as guests came and went. This would be more interesting, for it would imply that the rate of arrival of guests was exactly equal to the rate of departure. During a process of change, something is remaining constant. The significant conservation laws in nature are of this type, laws of constancy during change. It is not surprising that scientists, in their search for simplicity, fasten on conservation laws with particular enthusiasm, for what could be simpler than a quantity that remains absolutely constant during complicated processes of change. To cite an example from the world of particles, the total electric charge remains precisely constant in every collision, regardless of how many particles may be created or annihilated in the process.

The classical laws of physics are expressed primarily as laws of change, rather than as laws of constancy. Newton's law of motion describes how the motion of objects responds to forces that act upon them. Maxwell's equations of electromagnetism connect the rate of change of electric and magnetic fields in space and time. The early emphasis in fundamental science was rather naturally on discovering those laws which successfully describe the changes actually occurring in nature. Briefly, the "classical" philosophy con-

cerning nature's laws is this. Man can imagine countless possible laws, indeed infinitely many, that might describe a particular phenomenon. Of these, nature has chosen only one simple law, and the job of science is to find it. Having successfully found laws of change, man may derive from them certain conservation laws, such as the conservation of energy in mechanics. These appear as particularly interesting and useful consequences of the theory, but are not themselves taken as fundamental statements of the theory.

Gradually conservation laws have percolated to the top in the hierarchy of natural laws. This is not merely because of their simplicity, although this has been an important factor. It comes about also for two other reasons. One is the connection between conservation laws and principles of invariance and symmetry in nature—surely, one of the most beautiful aspects of modern science. The meaning of this connection will be discussed near the end of this chapter. The other reason we want to discuss here might best be described simply as a new view of the world, in which conservation laws appear naturally as the most fundamental statements of natural law. This new view is a view of order upon chaos—the order of conservation laws imposed upon the chaos of continual annihilation and creation taking place in the submicroscopic world. The strong hint emerging from recent studies of elementary particles is that the only inhibition imposed upon the chaotic flux of events in the world of the very small is that imposed by the conservation laws. Everything that *can* happen without violating a conservation law *does* happen.

This new view of democracy in nature—freedom under law—represents a revolutionary change in man's view of natural law. The older view of a fundamental law of nature was that it must be a law of *permission*. It defined what *can* (and must) happen in natural phenomena. According to the new view, the more fundamental law is a law of *prohibition*. It defines what *cannot* happen. A conservation law is, in effect, a law of prohibition. It prohibits any phenomenon that would change the conserved quantity, but otherwise allows any events. Consider, for example, the production of pions in a proton-proton collision,

$$p + p \rightarrow p + p + \pi + \pi + \pi + \cdots .$$

If a law of permission were operative, one might expect that, for protons colliding in a particular way, the law would specify the

number and the type of pions produced. A conservation law is less restrictive. The conservation of energy limits the number of pions that can be produced, because the mass of each one uses up some of the available energy. It might say, for example, that not more than six pions can be produced. In the actual collision there might be none, or one, or any number up to six. The law of charge conservation says that the total charge of the pions must be zero, but places no restriction on the charge of any particular pion; this could be positive, negative, or neutral.

To make more clear the distinction between laws of permission and laws of prohibition, let us return to the cocktail party. A law of change, which is a law of permission, might describe the rate of arrival and the rate of departure of guests as functions of time. In simplest form, it might say that three guests per minute arrive at 6:00, two guests per minute at 6:15, and so on. Or it might say, without changing its essential character as a law of permission, that the rate of arrival of guests is given by the formula:

$$R = \frac{A}{\pi D} \frac{1}{1 + \left(T - 5 - 2\dfrac{A}{D}\right)^2},$$

where R is the number of guests arriving per minute, A is the annual income of the host in thousands of dollars, D is the distance in miles from the nearest metropolitan center, and T is the time of day. This law resembles, in spirit, a classical law of physics. It covers many situations, but for any particular situation it predicts exactly what will happen.

A conservation law is simpler and less restrictive. Suppose it is observed that between 7 and 10 o'clock the number of guests is conserved at all parties. This is a grand general statement, appealing for its breadth of application and its simplicity. It would, were it true, be regarded as a deep truth, a very profound law of human behavior. But it gives much less detailed information than the formula for R above. The conservation law allows the guests to arrive at any rate whatever, so long as guests depart at the same rate. To push the analogy with natural law a bit further, we should say that according to the old view, since cocktail-party attendance is a fundamental aspect of human behavior, we seek and expect to find simple explicit laws governing the flow of guests. According to the new view, we expect to find the flux of arriving and depart-

ing guests limited only by certain conservation principles. Any behavior not prohibited by the conservation laws will, sooner or later, at some cocktail party, actually occur.

It should be clear that there is a close connection between this view of nature and the fundamental role of probability in nature. If the conservation law does not prohibit various possible results of an experiment, as in the proton-proton collision cited above, then these various possibilities will occur, each with some definite probability. The very fact that we can use the word chaos to describe the creation and annihilation events occurring continually among the particles rests on the existence of laws of probability. At best the probability, never the certainty, of these endless changes in the particle world can be known.

Are the laws of probability themselves derivable from conservation laws? The answer to this question is not yet known, but the trend of recent history is enough to make this author and many other physicists willing to bet on the affirmative. It appears possible, at least, that the conservation laws may not only be the most fundamental laws, but may be *all* the laws. They may be sufficient to characterize the elementary-particle world completely, specifying not only which events may occur and which are forbidden, but giving also the relative probabilities of those events which do occur.

We have so far emphasized that a conservation law is less restrictive than an explicit law of change, or law of permission. However, there are a number of different conservation laws and, taken all together, they may be very strongly restrictive, far more so than any one taken alone. In the ideal case, they may leave open only one possibility. The laws of prohibition, all taken together, then imply a unique law of permission. The most beautiful example of this kind of power of conservation laws concerns the nature of the photon. From conservation principles alone, it has been possible to show that the photon must be a massless particle of unit spin and no charge, emitted and absorbed by charged particles in a particular characteristic way. This truly amazing result has been expressed vividly by J. J. Sakurai who wrote recently, "The Creator was supremely imaginative when he declared, 'Let there be light.'"* In the world of human law, a man so hemmed in by restrictions that there is only one course of action open to him is

* *Annals of Physics*, Volume 11, page 5 (1960).

not very happy. In the world of natural law it is remarkable and satisfying to learn that a few simple statements about constant properties in nature can have locked within them such latent power that they determine uniquely the nature of light and its interaction with matter.

There are conservation laws and conservation laws. That is, some things in nature are constant, but others are even more constant. To convert this jargon into sense, some quantities in nature seem to be absolutely conserved, remaining unchanged in all events whatever; other quantities seem to be conserved in some kinds of processes and not in others. The rules governing the latter are still called conservation laws, but nature is permitted to violate them under certain circumstances. We shall postpone the discussion of these not-quite-conservation laws to Chapter Eight, and consider here only seven of the recognized absolute conservation laws. (There are two more absolute conservation laws of a more special kind, and they are also postponed to Chapter Eight.)

We begin by listing by name the seven quantities that are conserved:

1. Energy (including mass)
2. Momentum
3. Angular momentum, including spin
4. Charge
5. Electron-family number
6. Muon-family number
7. Baryon-family number.

There are two different kinds of quantities here, which can be called properties of motion and intrinsic properties, but the two are not clearly separated. The intrinsic particle properties that enter into the conservation laws are mass, spin, charge, and the several "family numbers." The properties of motion are kinetic energy, momentum, and angular momentum, the last frequently being called orbital angular momentum to avoid possible confusion with intrinsic spin, which is a form of angular momentum. In the laws of energy conservation and angular-momentum conservation, the intrinsic properties and properties of motion become mixed.

The interactions and transformations of the elementary particles serve admirably to illustrate the conservation laws and we shall

focus attention on the particles for illustrative purposes. It is through studies of the particles that all of these conservation laws have been verified, although the first four were already known in the macroscopic world. The particles provide the best possible testing ground for conservation laws, for any law satisfied by small numbers of particles is necessarily satisfied for all larger collections of particles, including the macroscopic objects of our everyday world. Whether the extrapolation of the submicroscopic conservation laws on into the cosmological domain is justified is uncertain, since gravity, whose effects in the particle world appear to be entirely negligible, becomes of dominant importance in the astronomical realm.

Various intrinsic properties of the particles were discussed in Chapter One, and we shall examine first the conservation laws that have to do with the intrinsic properties.

We learned in Chapter One that every particle carries the same electric charge as the electron (defined to be negative), or the equal and opposite charge of the proton (positive), or is neutral. The charge is a measure of the strength of electric force which the particle can exert and, correspondingly, a measure of the strength of electric force which the particle experiences. A neutral particle, of course, neither exerts nor responds to an electric force. A charged particle does both.

Using the proton charge as a unit, every particle's charge can be labeled $+1$, -1, or 0. The law of charge conservation requires that the total charge remain unchanged during every process of interaction or transformation. For any event involving particles, then, the total charge before the event must add up to the same value as the total charge after it. In the decay of a lambda into a neutron and a pion,

$$\Lambda^0 \rightarrow n + \pi^0,$$

the charge is zero both before and after. In the positive pion decay,

$$\pi^+ \rightarrow \mu^+ + \nu_\mu,$$

the products are a positive muon and a neutral neutrino. A possible high-energy nuclear collision might proceed as follows:

$$p + p \rightarrow n + \Lambda^0 + K^+ + \pi^+.$$

Neither positively charged proton survives the collision, but the net charge $+2$ appears on the particles created.

Notice that the law of charge conservation provides a partial explanation for the fact that particle charges come in only one size. If the charge on a pion were, say, 0.73 electron charges, it would be quite difficult to balance the books in transformation processes and maintain charge conservation. Actually, according to the present picture of elementary processes, the charge is conserved not only from "before" to "after," but at every intermediate stage of the process. One can visualize a single charge as an indivisible unit which, like a baton in a relay race, can be handed off from one particle to another, but never dropped or divided.

Perhaps the most salutary effect of the law of charge conservation in human affairs is the stabilization of the electron. The electron is the lightest charged particle and, for this reason alone, it cannot decay. The only lighter particles, the photon and neutrinos (and graviton), are neutral, and a decay of the electron would therefore necessarily violate the law of charge conservation. The stability of the electron is one of the simplest, yet one of the most stringent tests of the law of charge conservation. Nothing else prevents electron decay. If the law were almost, but not quite, valid, the electron should have a finite lifetime. A recent experiment places the electron lifetime beyond 10^{19} years; this means that charge conservation must be regarded as at least a very good approximation to an absolute law.

Unlike the other four laws, which were already known in the macroscopic world, the laws of family-number conservation were discovered through studies of particle transformation. We can best explain their meaning through examples. Recall that the proton and all heavier particles are called baryons, that is, they belong to the baryon family. In the decay of the unstable Λ particle,

$$\Lambda^0 \rightarrow p + \pi^-,$$

one baryon, the Λ, disappears, but another, the proton, appears. Similarly, in the decay of the Σ^0,

$$\Sigma^0 \rightarrow \Lambda^0 + \gamma,$$

the number of baryons is conserved. Notice that, in one of these examples, a pion is created, in the other, a photon. Pions and photons belong to none of the special family groups and can come and go

in any number. In a typical proton-proton collision the number of baryons (2) remains unchanged, as in the example,

$$p + p \rightarrow p + \Sigma^+ + K^0.$$

These and numerous other examples have made it appear that the number of baryons remains forever constant—in every single event, and therefore, of course, in any larger structure.

Each of the Ξ, Σ, and Λ particles, and the neutron, undergoes spontaneous decay into a lighter baryon. But the lightest baryon, the proton, has nowhere to go. The law of baryon conservation stabilizes the proton and makes possible the structure of nuclei and atoms and, therefore, of our world. From the particle physicist's point of view, this is a truly miraculous phenomenon, for the proton stands perched at a mass nearly 2,000 times the electron mass, having an intrinsic energy of about one billion electron volts, while beneath it lie the lighter unstable kaon, pion, and muon. Only the law of baryon conservation holds this enormous energy locked within the proton and makes it a suitable building block for the universe. The proton appears to be absolutely stable. If it is unstable it has, according to a recent experimental result, a half life greater than 7×10^{27} years, or about a billion billion times the age of the earth.

Our statement of the law of baryon conservation needs some amplification, for we have not yet taken into account antibaryons. A typical antiproton-production event at the Berkeley Bevatron might go as follows:

$$p + p \rightarrow p + p + p + \bar{p}.$$

(The bar over the letter designates the antiparticle. Since the antiproton has negative charge, the total charge of plus 2 is conserved.) It appears that we have transformed two baryons into four. Similarly, in the antiproton annihilation event,

$$p + \bar{p} \rightarrow \pi^+ + \pi^- + \pi^0,$$

two baryons have apparently vanished. The obvious way to patch up the law of baryon conservation is to assign to the antiparticles baryon number -1, and to the particles baryon number $+1$. Then the law would read: In every event the total number of baryons *minus* the total number of antibaryons is conserved; or, equivalently, the total baryon number remains unchanged.

The cynic might say that with so many arbitrary definitions—which particles should be called baryons and which not, and the use of negative baryon numbers—it is no wonder that a conservation law can be constructed. To this objection, two excellent answers can be given. The first is that it is not so easy to find an absolute conservation law. To find any quantity absolutely conserved in nature is so important that it easily justifies a few arbitrary definitions. The arbitrariness at this stage of history only reflects our lack of any deep understanding of the reason for baryon conservation, but it does not detract from the obvious significance of baryon conservation as a law of nature. The other answer, based on the mathematics of the quantum theory, is that the use of negative baryon number for antiparticles is perfectly natural, in fact, is demanded by the theory. This comes about because the description of the appearance of an antiparticle is "equivalent" (in a mathematical sense we cannot delve into) to the description of the disappearance of a particle; and conversely antiparticle annihilation is "equivalent" to particle creation.

The "electron family" contains only the electron and its neutrino, the "muon family" only the muon and its neutrino. For each of these small groups, there is a conservation of family members exactly like the conservation of baryons. The antiparticles must be considered negative members of the families, the particles positive members. These light-particle conservation laws are not nearly as well tested as the other absolute conservation laws because of the difficulties of studying neutrinos, but there are no known exceptions to them.

The beta decay of the neutron,

$$n \longrightarrow p + e^- + \overline{\nu_e},$$

illustrates nicely the conservation laws we have discussed. Initially, the single neutron has charge zero, baryon number 1, and electron-family number zero. The oppositely charged proton and electron preserve zero charge; the single proton preserves the baryon number; and the electron with its antineutrino ($\overline{\nu_e}$) together preserve zero electron-family number. In the pion decay processes,

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu \qquad \text{and} \qquad \pi^- \longrightarrow \mu^- + \overline{\nu_\mu},$$

muon-family conservation demands that a neutrino accompany the μ^+ antimuon, and an antineutrino accompany the μ^- muon. The muon, in turn, decays into three particles, for example,

$$\mu^- \rightarrow e^- + \nu_\mu + \overline{\nu_e},$$

which conserves the members of the muon family and of the electron family.

The general rule enunciated earlier in this chapter was that whatever *can* happen without violating a conservation law *does* happen. Until 1962, there was a notable exception to this rule; its resolution has beautifully strengthened the idea that conservation laws play a central role in the world of elementary particles. The decay of a muon into an electron and a photon,

$$\mu^- \rightarrow e^- + \gamma,$$

has never been seen, a circumstance which had come to be known as the μ-e-γ puzzle. Before the discovery of the muon's neutrino it was believed that electron, muon, and one neutrino formed a single family (called the lepton family) with a single family-conservation law. If this were the case, no conservation law prohibited the decay of muon into electron and photon, since the lost muon was replaced with an electron, and charge and all other quantities were conserved as well. According to the classical view of physical law, the absence of this process should have caused no concern. There was, after all, no law of permission which said that it should occur. There was only the double negative: No conservation law was known to prohibit the decay.

However, the view of the fundamental role of conservation laws in nature as the only inhibition on physical processes had become so ingrained in the thinking of physicists that the absence of this particular decay mode of the muon was regarded as a significant mystery. It was largely this mystery that stimulated the search for a second neutrino belonging exclusively to the muon. The discovery of the muon's neutrino established as a near certainty that the electron and muon belong to two different small families which are separately conserved. With the electron and muon governed by two separate laws of conservation, the prohibition of the μ-e-γ decay became immediately explicable, and the faith that what can happen does happen was further bolstered.

We turn now to the conservation laws which involve properties of motion.

In the world of particles there are only two kinds of energy: energy of motion, or kinetic energy, and energy of being, which is equivalent to mass. Whenever particles are created or annihilated (except the massless particles) energy is transformed from one form to the other, but the total energy in every process always remains conserved. The simplest consequence of energy conservation for the spontaneous decay of unstable particles is that the total mass of the products must be less than the mass of the parent. For each of the following decay processes the masses on the right add up to less than the mass on the left:

$$K^+ \to \pi^+ + \pi^+ + \pi^-,$$
$$\Xi^- \to \Lambda^0 + \pi^-,$$
$$\mu^+ \to e^+ + \nu_e + \overline{\nu_\mu}.$$

In particular, then, a massless particle cannot decay, and energy conservation prohibits every other "uphill" decay in which the products are heavier than the parent. An unstable particle at rest has only its energy of being, no energy of motion. The difference between this parent mass and the mass of the product particles is transformed into kinetic energy which the product particles carry away as they rapidly leave the scene.

One might suppose that if the parent particle is moving when it decays it has some energy of motion of its own which might be transformed to mass. The conservation of momentum prohibits this. The extra energy of motion is in fact "unavailable" for conversion into mass. If a particle loses energy, it also loses momentum. Momentum conservation therefore prohibits the conversion of all of the energy into mass. It turns out that momentum and energy conservation taken together forbid uphill decays into heavier particles no matter how fast the initial particle might be moving.

If two particles collide, on the other hand, some—but not all—of their energy of motion is available to create mass. It is in this way that the various unstable particles are manufactured in the laboratory. In an actual typical collision in the vicinity of an accelerator, one of the two particles, the projectile, is moving rapidly, and the other, the target, is at rest. Under these conditions, the requirement that the final particles should have as much momentum as the initial projectile severely restricts the amount of energy that

can be converted into mass. This is too bad, for the projectile has been given a great energy at a great expense. To make a proton-antiproton pair, for example, by the projectile-hitting-fixed-target method, the projectile must have a kinetic energy of 6 Bev (billion electron volts), of which only 2 Bev goes into making the mass. The 6 Bev Berkeley Bevatron was designed with this fact in mind in order to be able to make antiprotons and antineutrons. Typical processes for protons striking protons are:

$$p + p \rightarrow p + p + p + \bar{p},$$
$$p + p \rightarrow p + p + n + \bar{n}.$$

The unfortunate waste of 4 Bev in these processes could be avoided if the target proton were not quiescent, but flew at the projectile with equal and opposite speed. It is hard enough to produce one high-energy beam, and far more difficult to produce two at once. Nevertheless, the gain in available energy makes it worth the trouble, and a technique for producing "clashing beams" is now employed at Stanford University, where oppositely directed beams of electrons collide. The device is sometimes called by physicists the synchroclash.

Momentum is purely a property of motion—that is, if there is no motion, there is no momentum. It is somewhat trickier than energy, for momentum is what is called a vector quantity. It has direction as well as magnitude. Vectors are actually familiar in everyday life, whether or not we know them by that name. The velocity of an automobile is a vector, with a magnitude (50 miles per hour, for example) and a direction (northbound, for example). Force is a vector, a push or pull of some strength in some direction. Mass, on the other hand, is not a vector. It points in no particular direction. Energy also has no direction. The momentum of a rolling freight car, however, is directed along the tracks, and the momentum of an elementary particle is directed along its course through space.

In order to appreciate the law of momentum conservation, one must know how to add vectors. Two men pushing on a stalled car are engaged in adding vectors. If they push with equal strength *and* in the same direction, the total force exerted is twice the force each one exerts and, of course, in the direction they are pushing [Figure 4.1(a)]. If they push with equal strength but at opposite ends of the car, their effort comes to naught, for the sum of two vector quantities which are equal in strength but opposite in direction is

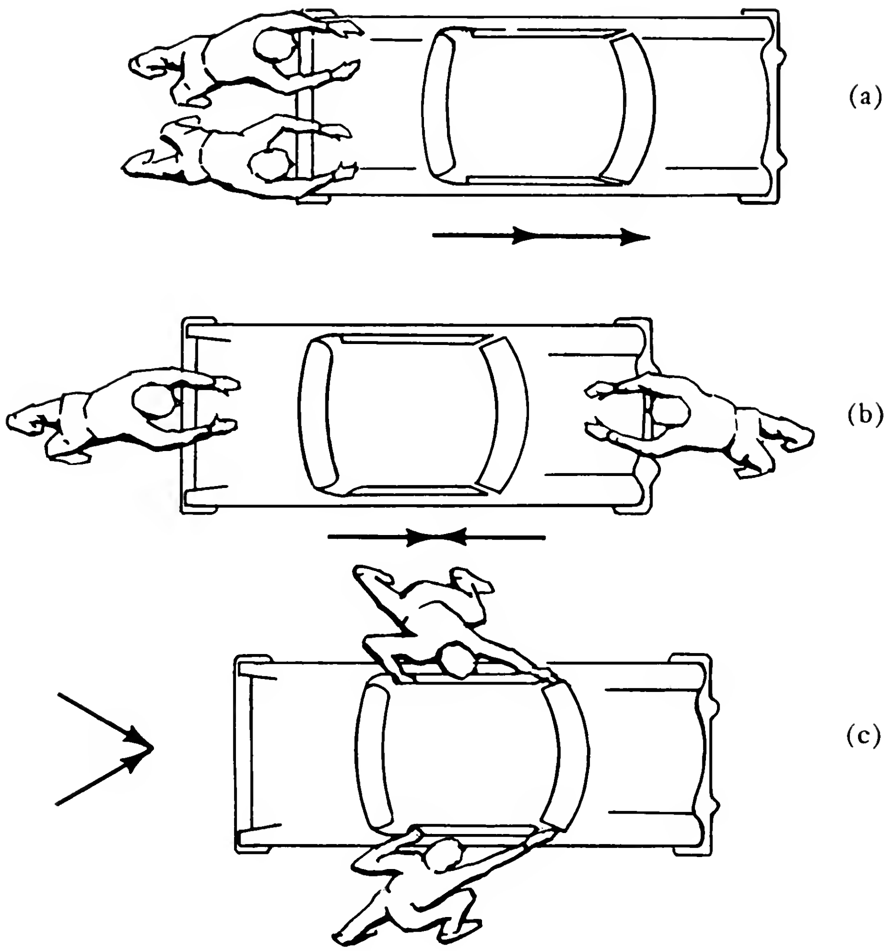


Figure 4.1. The addition of vectors. The forces exerted by two men pushing equally hard may be "added," that is, combined, to give any total from zero up to twice the force of each.

zero [Figure 4.1(b)]. If they get on opposite sides of the car and push partly inward, partly forward, the net force exerted will be forward, but less than twice the force of each [Figure 4.1(c)]. Depending on their degree of co-operation, the two men may achieve a strength of force from zero up to twice the force each can exert.

This is a general characteristic of the sum of two vectors. It may have a wide range of values depending on the orientation of the two vectors.

Consider the law of momentum conservation applied to the decay of a kaon into muon and neutrino,

$$K^+ \rightarrow \mu^+ + \nu_\mu.$$

Before the decay, suppose the kaon is at rest [Figure 4.2(a)]. After the decay, momentum conservation requires that muon and neutrino fly off with equal magnitudes of momenta *and* and that the momenta

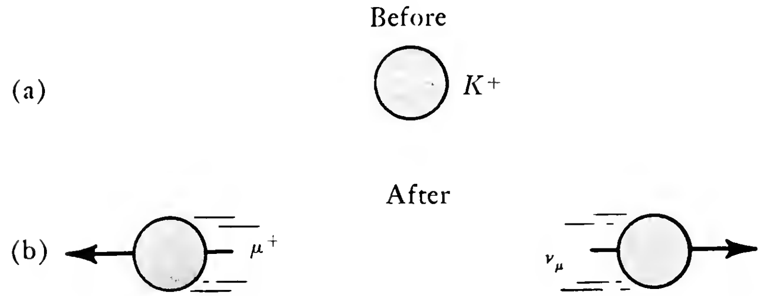


Figure 4.2. Momentum conservation in kaon decay. The total momentum is zero both before and after the decay.

be oppositely directed [Figure 4.2(b)]. Only in this way can the vector sum of the two final momenta be equal to the original momentum, namely zero. This type of decay, called a two-body decay, is rather common, and is always characterized by particles emerging in exactly opposite directions.

In a three-body decay, the emerging particles have more freedom. Figure 1.8, for example, shows the decay of a kaon into three pions with the tracks pointing in three different directions. Recalling the analogy between momentum and force, one can visualize a situation in which three different forces are acting and producing no net effect—two fighters and a referee all pushing in different directions in a clinch. Similarly, the momentum vectors must adjust themselves to produce no net effect; that is, they must add up

to give zero. Momentum conservation on a grander scale is shown in Figure 4.3, where eight particles emerge from a single event.

One vital prohibition of the law of momentum conservation is that against one-body decays. Consider, for example, this possibility,

$$K^+ \rightarrow \pi^+,$$

the transformation of kaon to pion. It satisfies the laws of charge and family-number conservation. It is consistent with energy conservation, for it is downhill in mass, and it also satisfies spin conservation. But the kaon-pion mass difference must get converted to energy of motion, so that if the kaon was at rest, the pion will fly away. In whatever direction it moves, it has some momentum and therefore violates momentum conservation, since the kaon had none. On the other hand, if we enforce the law of momentum conservation, and keep the pion at rest, we shall have violated energy conservation, for in this case the extra energy arising from the mass difference will be unaccounted for.

Angular momentum, a measure of the strength of rotational motion, has been a key concept in physics since the time of Kepler. Actually, Kepler did not recognize it as such, but the second of his three laws of planetary motion—the so-called law of areas—is equivalent to a law of conservation of angular momentum. According to this law, an imaginary straight line drawn from the earth to the sun sweeps out area in space at a constant rate. During a single day this line sweeps across a thin triangular region with apex at the sun and base along the earth's orbit. The area of this triangle is the same for every day of the year. So, when the earth is closer to the sun, it must move faster in order to define a triangle with the same area. It speeds up just enough, in fact, to maintain a constant value of its angular momentum, and the law of areas can be derived as a simple consequence of the law of conservation of angular momentum (this was first done by Newton).

The earth also serves to illustrate approximately the two kinds of angular momentum which enter into the conservation law—orbital and spin. The earth possesses angular momentum because of its orbital motion round the sun and because of its daily (spin) rotation about its own axis. For an elementary particle, the notion of spin is about the same as for the earth—rotational motion about an axis.

If a photographer in space took a time exposure of the earth and

Figure 4.3.

sun, his photograph would contain a short blur for the sun and a longer blur for the earth. He would notice that the blurs were not directed toward each other, and from this fact alone could conclude that earth and sun possess relative angular momentum. He would not need to know whether the earth swings around the sun or whether it proceeds into interstellar space. The key fact defining orbital angular momentum is some transverse motion of two objects. Any two moving objects, not aimed directly at each other, possess relative angular momentum. Two trains passing on the great plains have relative angular momentum, even though each is proceeding straight as an arrow. But if, through some mischance, both were on the same track on a collision course, they would have zero angular momentum. In particle collisions and decays, orbital angular momentum is usually of this trains-in-the-plains type, not involving actual orbiting of one particle round another. Figure 4.4 illustrates several examples of motion with angular momentum.

Angular momentum is a vector quantity. Its direction is taken to be the axis of rotation. The axis is well defined for spin, but what about orbital motion? For the passing trains, imagine again the blurred photograph indicating their direction of motion. Then ask: What would the axis be if the trains rotated about each other, instead of proceeding onward? The answer is a vertical axis; the angular momentum is directed upward. One more fact about orbital angular momentum needs to be known. Unlike spin, which comes in units of $\frac{1}{2}\hbar$, it comes only in units of $\hbar$.

The spinless pion decays into muon and neutrino, each with spin $\frac{1}{2}$. In Figure 4.5 we use artistic license and represent the particles by little spheres with arrows to indicate their direction of spin. Muon and neutrino spin oppositely in order to preserve the

---

Figure 4.3. Momentum conservation in an antiproton annihilation event. An antiproton entering from the bottom collides with a proton in the bubble chamber. Eight pions, four negative and four positive, spray off from the annihilation event in all directions. The momentum of each can be measured from the curvature of the track; the eight momenta added together as vectors are just equal to the momentum of the single incoming antiproton. (The kink in the track at the lower right is a pion decay, $\pi^+ \rightarrow \mu^+ + \nu_\mu$. In what general direction did the unseen neutrino fly off?)

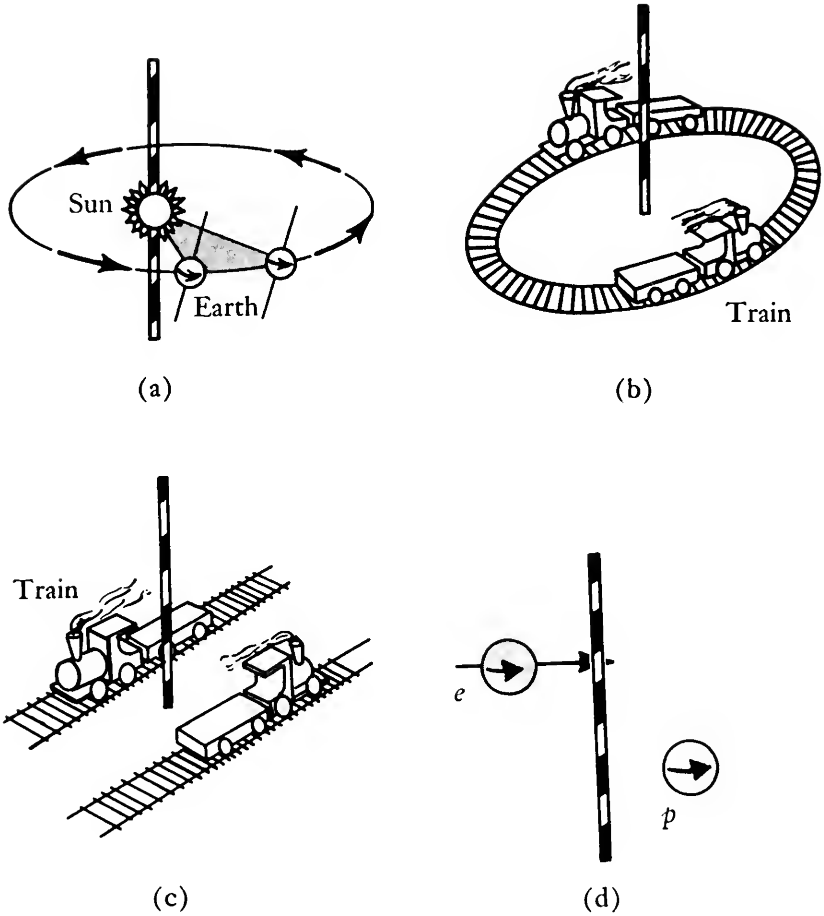


Figure 4.4. Examples of motion with angular momentum. (a) The earth possesses spin angular momentum about its axis as well as orbital angular momentum about an axis designated by the giant barber pole. The constancy of the earth's orbital angular momentum means that the shaded area swept out in one day is the same for every day of the year. (b) Trains on a circular track possess angular momentum about a vertical axis. (c) Even on straight tracks, a similar relative motion of trains represents angular momentum. (d) An electron flies past a proton. Both particles possess spin angular momentum and, because they are not on a collision course, they also have orbital angular momentum.

total zero angular momentum. In this case, no orbital angular momentum is involved.

Another two-body decay, that of the Λ, illustrates the coupling of spin and orbital motion. The Λ, supposed initially at rest [Figure 4.6(a)], has spin ½. One of its possible decay modes is

$$\Lambda^0 \rightarrow p + \pi^-.$$

This may proceed in two ways. The proton and pion may move apart with no orbital angular momentum, the proton spin directed upward to match the initial Λ spin [Figure 4.6(b)]; or the proton spin may be flipped to point downward while proton and pion

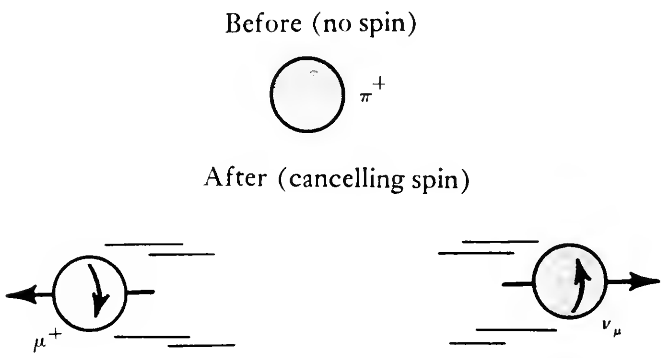


Figure 4.5. Angular-momentum conservation in pion decay. The total angular momentum is zero before and after the decay.

separate with one unit of orbital angular momentum, directed upward [Figure 4.6(c)]. In the first case,

original spin ½ (up) → final spin ½ (up).

In the second case,

original spin ½ (up) → final spin ½ (down) + orbital angular momentum 1 (up).

Beta decay, the earliest known particle decay process, serves nicely to illustrate all of the absolute conservation laws discussed. The beta decay of the neutron, indicated symbolically by

$$n \rightarrow p + e^- + \overline{\nu_e},$$

is pictured in Figure 4.7. Consider now the conservation laws applied to this decay.

Energy. Reference to Table 1 shows that the sum of the masses of the proton (1836.12), the electron (1.0), and the electron's

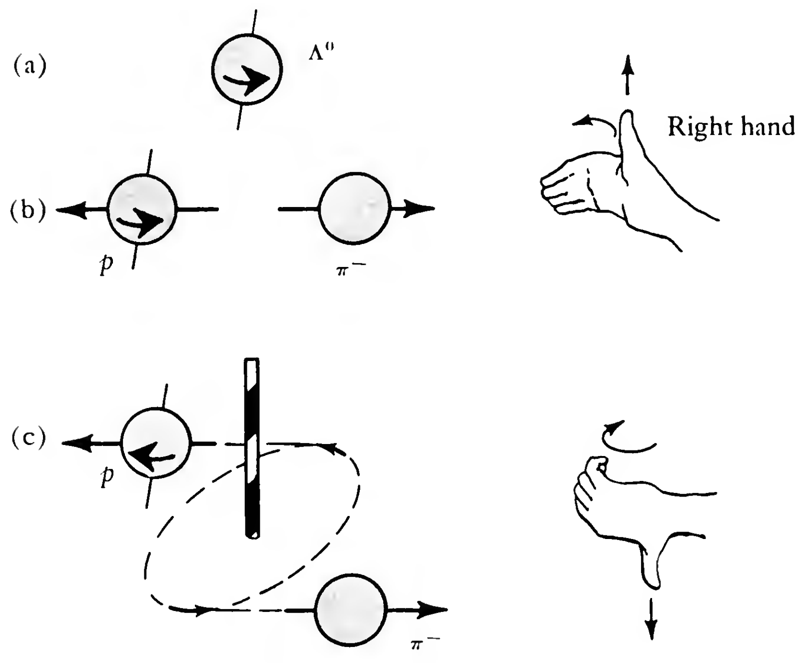


Figure 4.6. Angular-momentum conservation in lambda decay. The direction of angular momentum is defined by the right-hand rule. If the curved fingers of the right hand point in the direction of rotational motion, the right thumb defines the direction assigned to the angular momentum. Thus the particle spin is up in diagrams (a) and (b) and down in diagram (c); the orbital angular momentum is up in diagram (c).

neutrino (0), add up to less than the neutron mass (1838.65). The decay is therefore an allowed downhill decay, the slight excess mass going into kinetic energy of the products.

Momentum. The three particles must fan off in different direc-

tions with the available excess energy so distributed among them that the sum of the three momentum vectors is zero.

Angular momentum. One possibility, illustrated in Figure 4.7, is that the departing electron and proton have opposite cancelling spins, while the neutrino spins in the same direction as the original neutron to conserve the angular momentum.

Charge. The final charge (1 positive, 1 negative, 1 neutral) is zero, the same as the initial neutron charge.

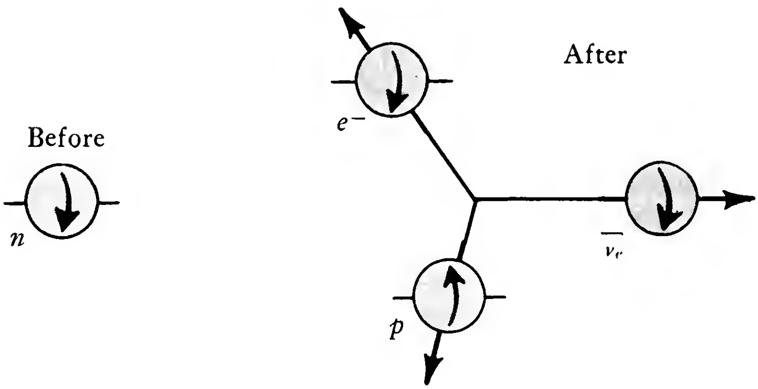


Figure 4.7. Beta decay of the neutron, $n \rightarrow p + e^- + \bar{\nu}_e$.

Electron-family number. The neutron has zero electron-family number. In the decay, one electron and one antineutrino ($\bar{\nu}_e$) are created to preserve zero electron-family number.

Muon-family number. No members of the muon family are created or destroyed.

Baryon number. The proton is the single baryon among the final three particles, matching the single original baryon.

Now we propose an exercise for the reader. Below are listed a few decays and transformations which do *not* occur in nature. If only one particle stands on the left, a decay process is understood. If two particles stand on the left, a collision process is understood. At least one conservation law prohibits each of these processes. Find at least one conservation law violated by each process. Several

violate more than one law and one of those listed violates five of the seven conservation laws.

a. $\mu^+ \rightarrow \pi^+ + \nu_\mu$
b. $e^- \rightarrow \nu_e + \gamma$
c. $p + p \rightarrow p + \Lambda^0 + \Sigma^+$
d. $\mu^+ \rightarrow \Lambda^0$
e. $n \rightarrow \mu^+ + e^- + \gamma$
f. $\Lambda^0 \rightarrow p + e^-$
g. $\pi^- + p \rightarrow \pi^- + n + \Lambda^0 + K^+$
h. $e^+ + e^- \rightarrow \mu^+ + \pi^-$
i. $\mu^- \rightarrow e^- + e^+ + \nu_\mu$

The aspect of conservation laws that makes them appear to the theorist and the philosopher to be the most beautiful and profound statements of natural law is their connection with principles of symmetry in nature. Baldly stated, energy, momentum, and angular momentum are all conserved because space and time are isotropic (the same in every direction) and homogeneous (the same at every place). This is a breath-taking statement when one reflects upon it, for it says that three of the seven absolute conservation laws arise solely because empty space has no distinguishing characteristics, and is everywhere equally empty and equally undistinguished. (Because of the relativistic link between space and time, we really mean space-time.) It seems, in the truest sense, that we are getting something from nothing.

Yet there can be no doubt about the connection between the properties of empty space and the fundamental conservation laws which govern elementary-particle behavior. This connection raises philosophical questions which we will mention but not pursue at any length. On the one hand, it may be interpreted to mean that conservation laws, being based on the most elementary and intuitive ideas, are the most profound statements of natural law. On the other hand, one may argue, as Bertrand Russell* has done, that it only demonstrates the hollowness of conservation laws ("truisms," according to Russell), energy, momentum, and angular momentum all being defined in just such a way that they must be conserved. Now, in fact, it is not inconsistent to hold both views at once. If

* Bertrand Russell, *The ABC of Relativity* (New York: New American Library, 1959).

the aim of science is the self-consistent description of natural phenomena based upon the simplest set of basic assumptions, what could be more satisfying than to have basic assumptions so completely elementary and self-evident (the uniformity of space-time) that the laws derived from them can be called truisms? Since the scientist generally is inclined to call most profound that which is most simple and most general, he is not above calling a truism profound. Speaking more pragmatically, we must recognize the discovery of *anything* that is absolutely conserved as something of an achievement, regardless of the arbitrariness of definition involved. Looking at those conservation laws whose basis we do not understand (the three family-number-conservation laws) also brings home the fact that it is easier to call a conservation law a truism after it is understood than before. It seems quite likely that we shall gain a deeper understanding of nature and of natural laws before the conservation of baryon number appears to anyone to be a self-evident truth.

Before trying to clarify through simple examples the connection between conservation laws and the uniformity of space, we consider the question, "What is symmetry?" In most general terms, symmetry means that when one thing (A) is changed in some particular way, something else (B) remains unchanged. A symmetrical face is one whose appearance (B) would remain the same if its two sides (A) were interchanged. If a square figure (A) is rotated through 90 degrees, its appearance (B) is not changed. Among plane figures, the circle is the most symmetrical, for if it is rotated about its center through any angle whatever, it remains indistinguishable from the original circle—or, in the language of modern physics, its form remains invariant. In the language of ancient Greece, the circle is the most perfect and most beautiful of plane figures.

Aristotle regarded the motion of the celestial bodies as necessarily circular because of the perfection (the symmetry) of the circle. Now, from a still deeper symmetry of space-time, we can derive the ellipses of Kepler. Modern science, which could begin only after breaking loose from the centuries-old hold of Aristotelian physics, now finds itself with an unexpected Aristotelian flavor, coming both from the increasingly dominant role of symmetry principles and from the increasingly geometrical basis of physics.

We are accustomed to think of symmetry in spatial terms. The

symmetry of the circle, the square, and the face are associated with rotations or inversions in space. Symmetry in time is an obvious extension of spatial symmetry; the fact that nature's laws appear to remain unchanged as time passes is a fundamental symmetry of nature. However, there exist some subtler symmetries, and it is reasonable to guess that the understanding of baryon conservation, for example, will come through the discovery of new symmetries not directly connected with space and time.

In the symmetry of interest to the scientist, the unchanging thing—the invariant element—is the form of natural laws. The thing changed may be orientation in space, or position in space or time, or some more abstract change (not necessarily realizable in practice) such as the interchange of two particles. The inversion of space and the reversal of the direction of flow of time are other examples of changes not realizable in practice, but nonetheless of interest for the symmetries of natural law. These latter two will be discussed in Chapter Eight.

If scientists in Chicago, New York, and Geneva perform the same experiment and get the same answer (within experimental error) they are demonstrating one of the symmetries of nature, the homogeneity of space. If the experiment is repeated later with the same result, no one is surprised, for we have come to accept the homogeneity of time. The laws of nature are the same, so far as we know, at all points in space, and for all times. This invariance is important and is related to the laws of conservation of energy and momentum, but ordinary experience conditions us to expect such invariance so that it seems at first to be trivial or self-evident. It might seem hard to visualize any science at all if natural law changed from place to place and time to time, but, in fact, quantitative science would be perfectly possible without the homogeneity of space-time. Imagine yourself, for example, on a merry-go-round that speeded up and slowed down according to a regular schedule. If you carried out experiments to deduce the laws of mechanics and had no way of knowing that you were on a rotating system, you would conclude that falling balls were governed by laws which varied with time and with position (distance from central axis), but you would be quite able to work out the laws in detail and predict accurately the results of future experiments, provided you knew where and when the experiment was to be carried out. Thanks to the actual homogeneity of space and time,

the results of future experiments can in fact be predicted without any knowledge of the where or when.

A slightly less obvious kind of invariance, although one also familiar from ordinary experience, is the invariance of the laws of nature for systems in uniform motion. Passengers on an ideally smooth train or in an ideally smooth elevator are unaware of motion. If the laws of mechanics were significantly altered, the riders would be aware of it through unusual bodily sensations. Such a qualitative guide is, of course, not entirely reliable, but careful experiments performed inside the ideal uniformly moving train would reveal the same laws of nature revealed by corresponding experiments conducted in a stationary laboratory. This particular invariance underlies the theory of relativity, and is a manifestation of the isotropy of four-dimensional space-time, a point we can regrettably not discuss in detail. What, to our limited three-dimensional vision, appears to be uniform motion is, to a more enlightened brain capable of encompassing four dimensions, merely a rotation. Instead of turning, say, from north to east, the experimenter who climbs aboard the train is, from the more general view, turning from space partly toward the time direction. According to relativity, which joins space and time together in a four-dimensional space-time, the laws of nature should no more be changed by "turning" experimental apparatus toward the time direction (that is, loading it aboard the train) than by turning it through 90 degrees in the laboratory.

The chain of connection we have been discussing is: Symmetry → invariance → conservation. The symmetry of space and time, or possibly some subtler symmetry of nature, implies the invariance of physical laws under certain changes associated with the symmetry. In the simplest case, for example, the symmetry of space which we call its homogeneity implies the invariance of experimental results when the apparatus is moved from one place to another. This invariance, in turn, implies the existence of certain conservation laws. The relation between conservation laws and symmetry principles is what we now wish to illuminate through two examples. Unfortunately, an adequate discussion of this important connection requires the use of mathematics beyond the scope of this book.

Suppose we imagine a single isolated hydrogen atom alone and at rest in empty space. If we could draw up a chair and observe

it without influencing it, what should we expect to see? (For this discussion, we ignore quantum mechanics and the wave nature of particles, pretending that electron and proton may be separately seen as particles, and be uninfluenced by the observer. The reader will have to accept the fact that these false assumptions are permissible and irrelevant for the present discussion.) We should see an electron in rapid motion circling about a proton, and the proton itself moving more slowly in a smaller circle. Were we to back off until the whole atom could only be discerned as a single spot, that spot, if initially motionless, would remain at rest forever. We now must ask whether this circumstance is significant or insignificant, important or dull. It certainly does not seem surprising. Why should the atom move, we may ask. It is isolated from the rest of the universe, no forces act upon it from outside, therefore there is nothing to set it into motion. If we leave a book on a table and come back later, we expect to find it there. Everyday experience conditions us to expect that an object on which no external forces act will not spontaneously set itself into motion. There is no more reason for the atom to begin to move than for the book to migrate across the table and fly into a corner. The trouble with this argument is that it makes use of the common sense of ordinary experience, without offering any explanation for the ordinary experience.

If we put aside "common sense" and ask what the atom might do, it is by no means obvious that it should remain at rest. In spite of the fact that no external forces are acting, strong internal forces are at work. The proton exerts a force on the electron which constantly alters its motion; the electron, in turn, exerts a force on the proton. Both atomic constituents are experiencing force. Why should these forces not combine to set the atom as a whole into motion? Having put the question in this way, we may consider the book on the table again. It consists of countless billions of atoms, each one exerting forces on its neighboring atoms. Through what miracle do these forces so precisely cancel out that no net force acts upon the book as a whole and it remains quiescent on the table?

The classical approach to this problem is to look for a positive, or permissive, law, a law which tells what *does* happen. Newton first enunciated this law which (except for some modification made necessary by the theory of relativity) has withstood the test of time to the present day. It is called Newton's third law, and says that all

forces in nature occur in equal and opposite balanced pairs. The proton's force on the electron is exactly equal and opposite to the electron's force on the proton. The sum of these two forces (the *vector* sum) is zero, so that there is no tendency for the structure as a whole to move in any direction. The balancing of forces, moreover, can be related to a balancing of momenta. By making use of Newton's second law,* which relates the motion to the force, one can discover that, in a hydrogen atom initially at rest, the balanced forces will cause the momenta of electron and proton to be equal and opposite. At a given instant, the two particles are moving in opposite directions. The heavier proton moves more slowly, but has the same momentum as the electron. As the electron swings to a new direction and a new speed in its track, the proton swings too in just such a way that its momentum remains equal and opposite to that of the electron. In spite of the continuously changing momenta of the two particles, the total momentum of the atom remains zero; the atom does not move. In this way—by "discovering" and applying two laws, Newton's second and third laws of motion—one derives the law of momentum conservation and finds an explanation of the fact that an isolated atom does not move.

Without difficulty, the same arguments may be applied to the book on the table. Since all forces come in equal and opposite pairs, the forces between every pair of atoms cancel, so that the total force is zero, no matter how many billions of billions of atoms and individual forces there might be.

It is worth reviewing the steps in the argument above. Two laws of permission were discovered, telling what does happen. One law relates the motion to the force; the other says that the forces between pairs of particles are always equal and opposite. From these laws, the conservation of momentum was derived as an interesting consequence, and this conservation law in turn explained the fact that an isolated atom at rest remains at rest.

The modern approach to the problem starts in quite a different way, by seeking a law of prohibition, a principle explaining why the atom does *not* move. This principle is the invariance of laws of nature to a change of position. Recall the chain of key ideas

---

* Newton's second law, usually written $F = ma$, says that the acceleration a experienced by a particle multiplied by its mass m is equal to the force F acting upon it. The law may also be stated in this way: The rate at which the momentum of a particle is changing is equal to the force applied.

referred to on page 105: symmetry → invariance → conservation. In the example of the isolated hydrogen atom, the symmetry of interest is the homogeneity of space. Founded upon this symmetry is the invariance principle just cited. Finally, the conservation law resting on this invariance principle is the conservation of momentum.

In order to clarify, through the example of the hydrogen atom, the connecting links between the assumed homogeneity of space and the conservation of momentum, we must begin with an exact statement of the invariance principle as applied to our isolated atom. The principle is this: No aspect of the motion of an isolated atom depends upon the location of the center of mass of the atom. The center of mass of any object is the average position of all of the mass in the object. In a hydrogen atom, the center of mass is a point in space between the electron and the proton, close to the more massive proton.

Let us visualize our hydrogen atom isolated in empty space with its center of mass at rest. Suppose now that its center of mass starts to move. In which direction should it move? We confront at once the question of the homogeneity of space. Investing our atom with human qualities for a moment, we can say that it has no basis upon which to "decide" how to move. To the atom surveying the possibilities, every direction is precisely as good or bad as every other direction. It is therefore frustrated in its "desire" to move and simply remains at rest.

This anthropomorphic description of the situation can be replaced by sound mathematics. What the mathematics shows is that an acceleration of the center of mass—for example, changing from a state of rest to a state of motion—is not consistent with the assumption that the laws of motion of the atom are independent of the location of the center of mass. If the center of mass of the atom is initially at rest at point *A* and it then begins to move, it will later pass through another point *B*. At point *A*, the center of mass had no velocity. At point *B* it does have a velocity. Therefore, the state of motion of the atom depends on the location of the center of mass, contrary to the invariance principle. Only if the center of mass remains at rest can the atom satisfy the invariance principle.* The immobility of the center of mass requires, in turn, that the two particles composing the atom have equal and opposite momenta.

* If the center of mass of the atom had been moving initially, the invariance principle requires that it continue moving with constant velocity.

A continual balancing of the two momenta means that their sum, the total momentum, is a constant.

The argument thus proceeds directly from the symmetry principle to the conservation law without making use of Newton's laws of motion. That this is a deeper approach to conservation laws, as well as a more esthetically pleasing one, has been verified by history. Although Newton's laws of motion have been altered by relativity and by quantum mechanics, the direct connection between the symmetry of space and the conservation of momentum has been unaffected—or even strengthened—by these modern theories and momentum conservation remains one of the pillars of modern physics. We must recognize that a violation of the law of momentum conservation would imply an inhomogeneity of space; this is not an impossibility, but it would have far-reaching consequences for our view of the universe.

Returning finally to the book on the table, we want to emphasize that the quiescence of the undisturbed book—a macroscopic object—at least strongly suggests that momentum conservation must be a valid law in the microscopic world. Viewed microscopically, the book is a collection of an enormous number of atoms, each one in motion. That this continuous microscopic motion never makes itself felt as spontaneous bulk motion of the whole book is true only because of the conservation of momentum which requires that every time an atom changes its momentum (as it is constantly doing) one or more other atoms must undergo exactly compensating changes of their momentum.

Through similar examples it is possible to relate the law of conservation of angular momentum to the isotropy of space. A compass needle which is held pointing east and is then released will swing toward the north because of the action of the earth's magnetic field upon it. But if the same compass needle is taken to the depths of empty space, far removed from all external influences, and set to point in some direction, it will remain pointing in that direction. A swing in one direction or the other would imply a nonuniformity* of space. If the uniformity of space is adopted as a fundamental symmetry principle, it can be concluded that the

* Strictly, momentum conservation rests on the *homogeneity* of space (uniformity of place), and angular momentum conservation rests on the *isotropy* of space (uniformity of direction). The distinction is not important for our purposes, and it is satisfactory to think of space simply as everywhere the same, homogeneity and isotropy being summarized by the word uniformity.

total angular momentum of all the atomic constituents of the needle must be a constant. Otherwise, the internal motions within the needle could set the whole needle into spontaneous rotation and its motion would violate the symmetry principle.

Energy conservation, in a way that is not so easy to see, is related to the homogeneity of time. Thus all three conservation laws—of energy, momentum, and angular momentum—are "understood" in terms of the symmetry of space-time, and indeed the theory of relativity has shown that these three laws are all parts of a single general conservation law in the four-dimensional world.

Only one of the three conservation laws governing the intrinsic properties of the particles has so far been understood in terms of a symmetry principle. This is the law of charge conservation. (Recall, however, that the *quantization* of charge is not yet understood.) The symmetry principle underlying charge conservation is considerably more subtle than the space-time symmetry underlying the conservation laws of properties of motion. The modern version of this symmetry principle rests upon technical aspects of the theory of quantum mechanics (it may be based also on equally technical aspects of the theory of electromagnetism). Nevertheless, it is such a stunning victory for the power of a symmetry principle that we must try, however crudely, to indicate the modern view of this symmetry.

In the main, the classical theories of physics deal directly with quantities which are measurable, usually called observables. Force, mass, velocity, and almost all the other concepts described by the classical laws are themselves observables. The equations of quantum mechanics, however, contain quantities which are not themselves observables. From these quantities—one step removed from reality—the observables are derived. The "wave function" is one of the unobservable quantities; it determines the probability, say, that the electron is at any particular point in the hydrogen atom, but is itself not that probability nor any other measurable thing. Now enters the idea of symmetry. Any change that can be made in the unobservable quantity without resulting in a change of the observables ought to leave all the laws of nature unchanged. After careful scrutiny, this statement seems so obviously true that it is hard to understand how it could have any important consequences. Of course one ought to be able to do anything whatever to unobservable quantities so long as observables are not changed. But remember

how important were the properties of empty space. Equally important are the properties of unobservables such as wave functions.

Space itself may be regarded as an unobservable. The uniformity of space means that it is impossible, by any experimental means, to ascertain one's absolute position in space. An experiment carried out at one place will yield results identical to the results of the same experiment carried out at another place. Any change in the unobservable space (for instance, moving the apparatus from one place to another) must leave unchanged the laws of nature and the observable results of experiment. As we have just seen, this symmetry principle or invariance requirement underlies the law of momentum conservation.

When an analogous symmetry principle is applied to the unobservable wave function of the electron a conservation law results, the conservation of charge. Expressed negatively, if charge were not conserved, the form of the equations of quantum mechanics would depend upon unobservable quantities, a situation at variance with our symmetry principle. The analogous statement for spatial homogeneity would be: If momentum were not conserved, the laws of mechanics would depend upon the absolute location in space and such dependence is at variance with the assumed symmetry of space.

Regrettably, we can not explain the law of charge conservation more fully without mathematics. It is expected, but not yet verified, that some undiscovered subtle symmetries of nature underlie the laws of electron-family conservation, muon-family conservation, and baryon conservation. The absolute prohibition of proton decay, which keeps its enormous intrinsic energy locked forever in the form of mass, can be no accident, but the reason still remains hidden.

Answers

The particle transformations listed on page 102 *violate* the following conservation laws:

a. Energy (an "uphill" decay); muon-family number (since μ^+ is an antiparticle).
b. Charge.
c. Angular momentum; baryon number.
d. Energy; momentum (a one-particle decay); charge; muon-

family number; baryon number.

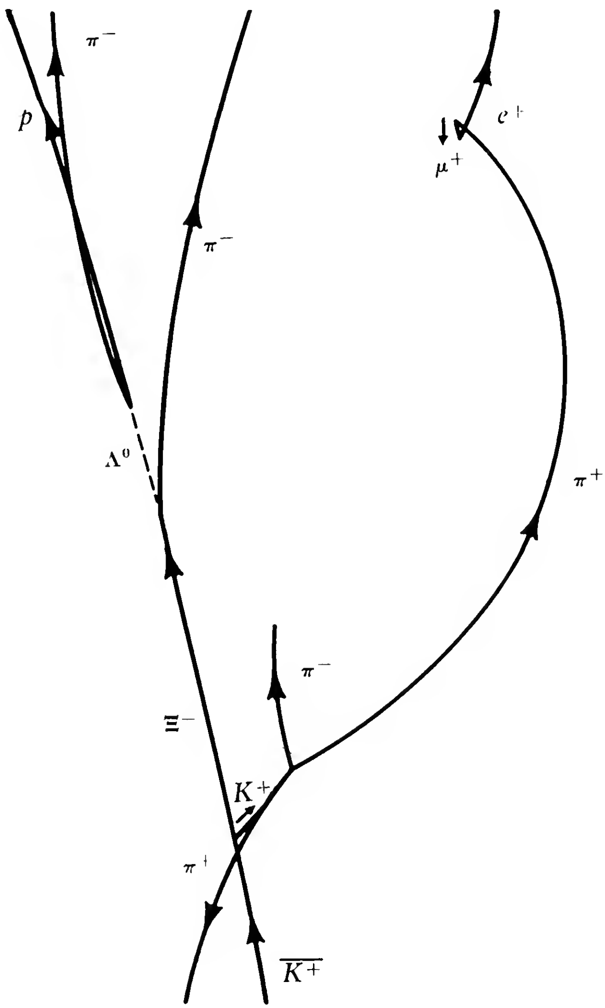
e. Angular momentum; baryon number; muon-family number; electron-family number.

f. Angular momentum; electron-family number.

g. Angular momentum; baryon number.

h. Angular momentum; muon-family number.

i. Charge. (Why is angular momentum conservation satisfied?)



Schematic analysis of the photograph on the opposite page.

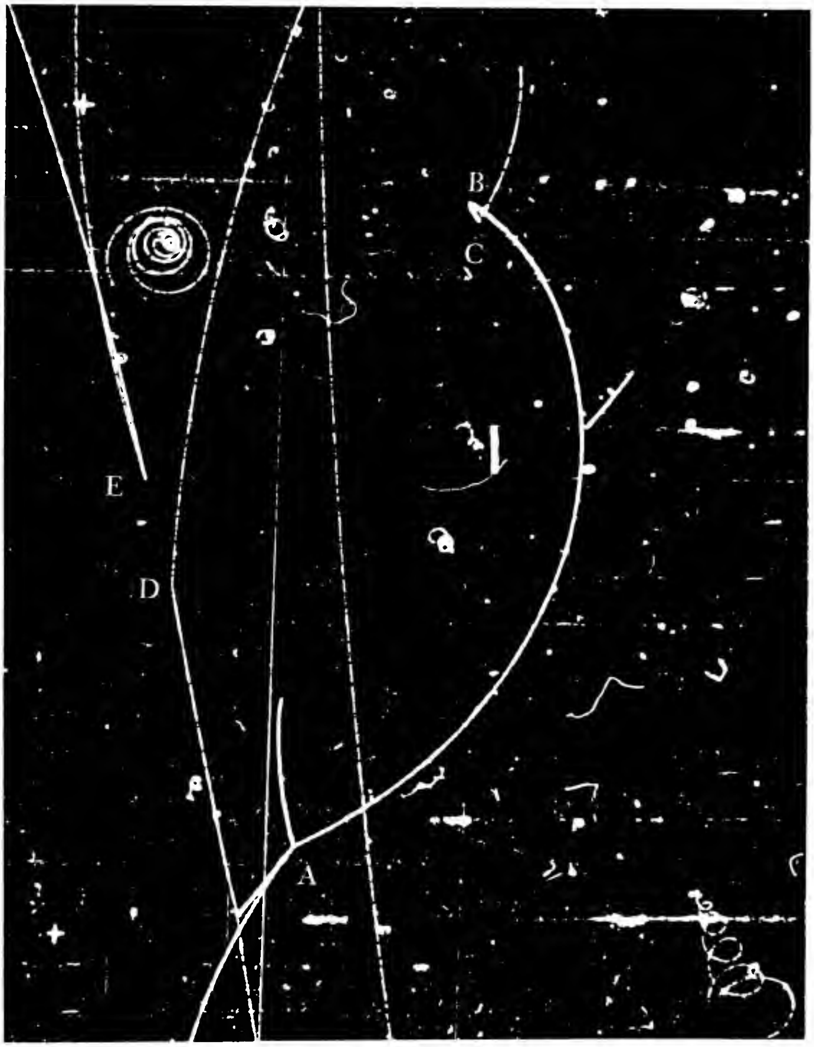


Figure 1.8. Decay of unstable particles. This unusual bubble-chamber photograph shows the decay of five different elementary particles. At point *A*, a positive kaon decays into three pions. At *B*, one of these pions decays into a muon and an unseen neutrino. At *C*, the muon decays into a positron (plus two neutrinos). At point *D*, a xi particle decays into a lambda particle and a pion. The invisible neutral lambda decays into a proton and a pion at point *E*.

Until 1956 the laws of physics included no preference for "right-handedness" or "left-handedness." But in 1956 the "law of parity" failed in experiments involving elementary particles, indicating that the universe is in some sense asymmetric.

---

17 The Fall of Parity

Martin Gardner

Chapter from his book, *The Ambidextrous Universe,* published in 1964.

As far as anyone knows at present, all events that take place in the universe are governed by four fundamental types of forces (physicists prefer to say "interactions" instead of "forces," but there is no harm in using here the more common term):

1. Nuclear force.
2. Electromagnetic force.
3. Weak interaction force.
4. Gravitational force.

The forces are listed in decreasing order of strength. The strongest, nuclear force, is the force that holds together the protons and neutrons in the nucleus of an atom. It is often called the "binding energy" of the nucleus. Electromagnetism is the force that binds electrons to the nucleus, atoms into molecules, molecules into liquids and solids. Gravity, as we all know, is the force with which one mass attracts another mass; it is the force chiefly responsible for binding together the substances that make up the earth. Gravitational force is so weak that unless a mass is enormously large it is extremely difficult to measure. On the level of the elementary particles its influence is negligible.

The remaining force, the force involved in "weak interactions," is the force about which the least is known. That such a force must exist is indicated by the fact that in certain decay interactions involving particles (such as beta-decay, in which electrons or positrons are shot out from radioactive nuclei), the speed of the reaction is much slower than it would be if either nuclear or electromagnetic forces were responsible. By "slow" is meant a reaction of, say, one ten-billionth of a second. To a nuclear physicist this is an exceedingly lazy effect—about a ten-trillionth the speed of reactions in which nuclear force is involved. To explain this lethargy it has been necessary to assume a force weaker than electromagnetism but stronger than the extremely weak force of gravity.

The "theta tau puzzle," over which physicists scratched their heads in 1956, arose in connection with a weak interaction involving a "strange particle" called the K-meson. (Strange particles are a class of recently discovered particles called "strange" because they do not seem to fit in anywhere with any of the other particles.) There appeared to be two distinct types of K-mesons. One, called the theta meson, decayed into two pi mesons. The other, called the tau meson, decayed into three pi mesons. Nevertheless, the two types of K-mesons seemed to be indistinguishable from each other. They had precisely the same mass, same charge, same lifetime. Physicists would have liked to say that there was only one K-meson; sometimes it decayed into two, sometimes into three pi mesons. Why didn't they? Because it would have meant that parity was not conserved. The theta meson had even parity. A pi meson has odd parity. Two pi mesons have a total parity that is even, so parity is conserved in the decay of the theta meson. But *three* pi mesons have a total parity that is odd.

Physicists faced a perplexing dilemma with the following horns:

1. They could assume that the two K-mesons, even though indistinguishable in properties, were really two different par-

ticles: the theta meson with even parity, the tau meson with odd parity.

2. They could assume that in one of the decay reactions parity was not conserved.

To most physicists in 1956 the second horn was almost unthinkable. As we saw in Chapter 20, it would have meant admitting that the left–right symmetry of nature was being violated; that nature was showing a bias for one type of handedness. The conservation of parity had been well established in all "strong" interactions (that is, in the nuclear and electromagnetic interactions). It had been a fruitful concept in quantum mechanics for thirty years.

In April, 1956, during a conference on nuclear physics at the University of Rochester, in New York, there was a spirited discussion of the theta-tau puzzle. Richard Phillips Feynman,[1] a physicist at the California Institute of Technology, raised the question: Is the law of parity sometimes violated? In corresponding with Feynman, he has given me some of the details behind this historic question. They are worth putting on record.

The question had been suggested to Feynman the night before by Martin Block, an experimental physicist with whom Feynman was sharing a hotel room. The answer to the theta-tau puzzle, said Block, might be very simple. Perhaps the lovely law of parity does not always hold. Feynman responded by pointing out that if this were true, there would be a way to distinguish left from right. It would be surprising, Feynman said, but he could think of no way such a notion conflicted with known experimental results. He promised Block he would raise the question at next day's meeting to see if anyone could find anything wrong with the idea. This he did, prefacing his remarks with, "I am asking this question for Martin Block." He regarded the notion as such an interesting one that, if it turned out to be true, he wanted Block to get credit for it.

Chen Ning Yang and his friend Tsung Dao Lee, two young and brilliant Chinese-born physicists, were present at the meet-

ing. One of them gave a lengthy reply to Feynman's question.

"What did he say?" Block asked Feynman later.

"I don't know," replied Feynman. "I couldn't understand it."

"People teased me later," writes Feynman, "and said my prefacing remark about Martin Block was made because I was afraid to be associated with such a wild idea. I thought the idea unlikely, but possible, and a very exciting possibility. Some months later an experimenter, Norman Ramsey, asked me if I believed it worth while for him to do an experiment to test whether parity is violated in beta decay. I said definitely yes, for although I felt sure that parity would *not* be violated, there was a possibility it would be, and it was important to find out. 'Would you bet a hundred dollars against a dollar that parity is not violated?' he asked. 'No. But fifty dollars I will.' 'That's good enough for me. I'll take your bet and do the experiment.' Unfortunately, Ramsey didn't find time to do it then, but my fifty dollar check may have compensated him slightly for a lost opportunity."

During the summer of 1956 Lee and Yang thought some more about the matter. Early in May, when they were sitting in the White Rose Cafe near the corner of Broadway and 125th Street, in the vicinity of Columbia University, it suddenly struck them that it might be profitable to make a careful study of all known experiments involving weak interactions. For several weeks they did this. To their astonishment they found that although the evidence for conservation of parity was strong in all strong interactions, there was no evidence at all for it in the weak. Moreover, they thought of several definitive tests, involving weak interactions, which would settle the question one way or the other. The outcome of this work was their now-classic paper "Question of Parity Conservation in Weak Interactions."

"To decide unequivocally whether parity is conserved in weak interactions," they declared, "one must perform an experiment to determine whether weak interactions differentiate the

right from the left. Some such possible experiments will be discussed."

Publication of this paper in *The Physical Review* (October 1, 1956) aroused only mild interest among nuclear pysicists. It seemed so unlikely that parity would be violated that most physicists took the attitude: Let someone else make the tests. Freeman J. Dyson, a physicist now at the Institute for Advanced Study in Princeton, writing on "Innovation in Physics" (*Scientific American*, September 1958) had these honest words to say about what he called the "blindness" of most of his colleagues.

"A copy of it [the Lee and Yang paper] was sent to me and I read it. I read it twice. I said, 'This is very interesting,' or words to that effect. But I had not the imagination to say, 'By golly, if this is true it opens up a whole new branch of physics.' And I think other physicists, with very few exceptions, at that time were as unimaginative as I."

Several physicists were prodded into action by the suggestions of Lee and Yang. The first to take up the gauntlet was Madame Chien-Shiung Wu, a professor of physics at Columbia University and widely regarded as the world's leading woman physicist. She was already famous for her work on weak interactions and for the care and elegance with which her experiments were always designed. Like her friends Yang and Lee, she, too, had been born in China and had come to the United States to continue her career.

The experiment planned by Madame Wu involved the beta-decay of cobalt-60, a highly radioactive isotope of cobalt which continually emits electrons. In the Bohr model of the atom, a nucleus of cobalt 60 may be thought of as a tiny sphere which spins like a top on an axis labeled north and south at the ends to indicate the magnetic poles. The beta-particles (electrons) emitted in the weak interaction of beta-decay are shot out from both the north and the south ends of nuclei. Normally, the nuclei point in all directions, so the electrons are shot out

in all directions. But when cobalt-60 is cooled to near absolute zero (−273 degrees on the centigrade scale), to reduce all the joggling of its molecules caused by heat, it is possible to apply a powerful electromagnetic field which will induce more than half of the nuclei to line up with their north ends pointing in the same direction. The nuclei go right on shooting out electrons. Instead of being scattered in all directions, however, the electrons are now concentrated in two directions: the direction toward which the north ends of the magnetic axes are pointing, and the direction toward which the south ends are pointing. If the law of parity is not violated, there will be just as many electrons going one way as the other.

To cool the cobalt to near absolute zero, Madame Wu needed the facilities of the National Bureau of Standards, in Washington, D. C. It was there that she and her colleagues began their historic experiment. If the number of electrons divided evenly into two sets, those that shot north and those that shot south, parity would be preserved. The theta-tau puzzle would remain puzzling. If the beta-decay process showed a handedness, a larger number of electrons emitted in one direction than the other, parity would be dead. A revolutionary new era in quantum theory would be under way.

At Zurich, one of the world's greatest theoretical physicists, Wolfgang Pauli, eagerly awaited results of the test. In a now famous letter to one of his former pupils, Victor Frederick Weisskopf (then at the Massachusetts Institute of Technology), Pauli wrote: "I do *not* believe that the Lord is a weak left-hander, and I am ready to bet a very high sum that the experiments will give symmetric results."

Whether Pauli (who died in 1958) actually made (like Feynman) such a bet is not known. If he did, he also lost. The electrons in Madam Wu's experiment were *not* emitted equally in both directions. Most of them were flung out from the south end; that is, the end toward which a majority of the cobalt-60 nuclei pointed their south poles.

At the risk of being repetitious, and possibly boring readers who see at once the full implication of this result, let us pause to make sure we understand exactly why Madam Wu's experiment is so revolutionary. It is true that the *picture* (Figure 62)

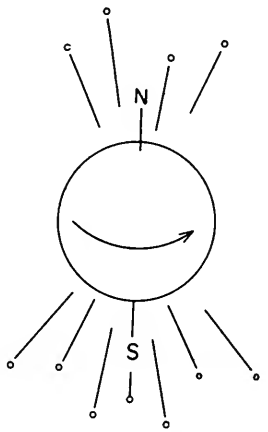


FIGURE 62. An electron is more likely to be flung out from the south end of a cobalt-60 nucleus than from its north end.

of the cobalt-60 nucleus, spinning in a certain direction around an axis labeled *N* and *S*, is an asymmetric structure not superposable on its mirror image. But this is just a picture. As we have learned, the labeling of *N* and *S* is purely conventional. There is nothing to prevent one from switching *N* and *S* on all the magnetic fields in the universe. The north ends of cobalt-60 nuclei would become south, the south ends north, and a similar exchange of poles would occur in the electromagnetic field used for lining up the nuclei. Everything prior to Madame Wu's experiment suggested that such a switch of poles would not make a measurable change in the experimental situation. If there were some intrinsic, observable difference between poles—one red and one green, or one strong and one weak—then the labeling of *N* and *S* would be more than a convention. The

cobalt-60 nuclei would possess true spatial asymmetry. But physicists knew of no way to distinguish between the poles except by testing their reaction to other magnetic axes. In fact, as we have learned, the poles do not really exist. They are just names for the opposite sides of a spin.

Madame Wu's experiment provided for the first time in the history of science a method of labeling the ends of a magnetic axis in a way that is not at all conventional. *The south end is the end of a cobalt-60 nucleus that is most likely to fling out an electron!*

The nucleus can no longer be thought of as analogous to a spinning sphere or cylinder. It must now be thought of as analogous to a spinning cone. Of course, this is no more than a metaphor. No one has the slightest notion at the moment of why or how one end of the axis is different, in any intrinsic way, from the other. But there is a difference! "We are no longer trying to handle screws in the dark with heavy gloves," was the way Sheldon Penman of the University of Chicago put it (*Scientific American*, July 1961), "we are being handed the screws neatly aligned on a tray, with a little searchlight on each that indicates the direction of its head."

It should be obvious now that here at long last is a solution to the Ozma problem—an experimental method of extracting from nature an unambiguous definition of left and right. We say to the scientists of Planet X: "Cool the atoms of cobalt-60 to near absolute zero. Line up their nuclear axes with a powerful magnetic field. Count the number of electrons flung out by the two ends of the axes. The end that flings out the most electrons is the end that we call 'south.' It is now possible to label the ends of the magnetic axis of the field used for lining up the nuclei, and this in turn can be used for labeling the ends of a magnetic needle. Put such a needle above a wire in which the current moves away from you. The north pole of this needle will point in the direction we call 'left.'"

We have communicated precisely and unambiguously to

Planet X our meaning of the word 'left.' Neither we nor they will be observing in common any single, particular asymmetric structure. We will be observing in common a universal law of nature. In the weak interactions, nature herself, by her own intrinsic handedness, has provided an operational definition of left and right! It is easy to understand why Pauli and other physicists did not expect Madame Wu's experiment to overthrow parity. It would have meant that nature is not ambidextrous!

In the context of my *Esquire* tale about left and right, the cobalt-60 experiment provides a method by which the puzzled astronauts could tell whether they were reversed. Of course they would have to find some cobalt on the unknown planet, convert it to its radioactive isotope by bombarding it with neutrons, and so on. But assuming that they had the equipment and could find the necessary materials, they would be able to test their handedness.

Similarly, Madame Wu's experiment clearly violates the assertion that all natural events can be photographed on motion picture film and projected in reversed form without the viewer being the wiser.

EXERCISE 16: *Explain precisely how an observation of all details of the cobalt-60 experiment, when viewed as a projected motion picture, would enable one to tell whether the film had been reversed.*

Athough evidence against the conservation of parity was strongly indicated by Madame Wu's work in late 1956, the experiment was not finally completed until early in January 1957. Results were formally announced by Columbia University's distinguished physicist Isador Rabi on January 15, 1957. The announcement also included the results of a confirming experiment conducted by Columbia physicists at the Nevis Cyclotron Laboratories at Irvington-on-Hudson in Westchester

County, New York. This confirming test, made with mu mesons, showed an even stronger handedness. The mu mesons shot out twice as many electrons in one direction as in the other. Independent of both experiments, a third test was made at the University of Chicago using the decay of pi and mu mesons. It, too, showed violation of parity. All over the world physicists began testing parity in other weak interactions. By 1958 it was apparent that parity is violated in *all* such interactions. The theta-tau puzzle was solved. There is only *one* K-meson. Parity is *not* conserved.

"A rather complete theoretical structure has been shattered at the base," declared Rabi (quoted by the *New York Times*, January 16, 1957), "and we are not sure how the pieces will be put together." An unnamed physicist was reported by the *Times* as saying that nuclear physics had been battering for years at a closed door only to discover suddenly that it wasn't a door at all—just a picture of a door painted on a wall. Now, he continued, we are free to look around for the true door. O. R. Frisch, the physicist who was a co-discoverer of nuclear fission, reports in his book *Atomic Physics Today* (Basic, 1961) that on January 16, 1957, he received the following air letter from a friend:

> Dear Robert:
>
> HOT NEWS. Parity is not conserved. Here in Princeton they talk about nothing else; they say it is the most important result since the Michelson experiment . . .

The Michelson experiment was the famous Michelson-Morley test in 1887 which established the constant velocity of light regardless of the motion of source and observer—a historic experiment which paved the way for Einstein's theory of relativity. Madame Wu's experiment may well prove to be equally historic.

The two tests were very much alike in their shattering element of surprise. Everybody expected Albert Michelson and Edward

Morley to detect a motion of the earth relative to a fixed "ether." It was the negative result of this test that was so upsetting. Everybody expected Madame Wu to find a left–right symmetry in the process of beta-decay. Nature sprang another surprise! It was surprising enough that certain particles had a handedness; it was more surprising that handedness seemed to be observable only in weak interactions. Physicists felt a shock even greater than Mach had felt when he first encountered the needle-and-wire asymmetry.

"Now after the first shock is over," Pauli wrote to Weisskopf on January 27, after the staggering news had reached him, "I begin to collect myself. Yes, it was very dramatic. On Monday, the twenty-first, at 8 P.M. I was supposed to give a lecture on the neutrino theory. At 5 P.M. I received three experimental papers [reports on the first three tests of parity]. . . . I am shocked not so much by the fact that the Lord prefers the left hand as by the fact that he still appears to be left-handed symmetric when he expresses himself strongly. In short, the actual problem now seems to be the question: Why are strong interactions right-and-left symmetric?

The Indian physicist Abdus Salam (from whose article on "Elementary Particles" in *Endeavor*, April 1958, the extracts from Pauli's letters are taken) tried to explain to a liberal-arts-trained friend why the physicists were so excited about the fall of parity. "I asked him," wrote Salam in this article, "if any classical writer had ever considered giants with only the left eye. He confessed that one-eyed giants have been described, and he supplied me with a full list of them; but they always sport their solitary eye in the middle of the forehead. In my view, what we have found is that space is a weak left-eyed giant."

Physicist Jeremy Bernstein, in an article on "A Question of Parity" which appeared in *The New Yorker*, May 12, 1962, reveals an ironic sidelight on the story of parity's downfall. In 1928 three physicists at New York University had actually discovered a parity violation in the decay of a radioactive isotope of radium! The experiment had been repeated with refined

techniques in 1930. "Not only in every run," the experimenter reported, "but even in all readings in every run, with few exceptions," the effect was observable. But this was at a time when, as Bernstein puts it, there was no theoretical context in which to place these results. They were quickly forgotten. "They were," writes Bernstein, "a kind of statement made in a void. It took almost thirty years of intensive research in all branches of experimental and theoretical physics, and, above all, it took the work of Lee and Yang, to enable physicists to appreciate exactly what those early experiments implied."

In 1957 Lee and Yang received the Nobel prize in physics for their work. Lee was then 30, Yang 34. The choice was inevitable. The year 1957 had been the most stirring in modern particle physics, and Lee and Yang had done most of the stirring. Today the two men have adjacent offices at the Institute for Advanced Study in Princeton, where they continue to collaborate. Both live in Princeton with their attractive wives and children, proud of their Chinese heritage, deeply committed to science, and with a wide range of interests outside of physics and mathematics. If you are curious to know more about these two remarkable men, look up Bernstein's excellent *New Yorker* article.

It is worth pausing to note that, like so many other revolutions in physics, this one came about as the result of largely abstract, theoretical, mathematical work. Not one of the three experiments that first toppled parity would have been performed at the time it was performed if Lee and Yang had not told the experimenters what to do. Lee had had no experience whatever in a laboratory. Yang had worked briefly in a lab at the University of Chicago, where he was once a kind of assistant to the great Italian physicist Enrico Fermi. He had not been happy in experimental work. His associates had even made up a short rhyme about him which Bernstein repeats:

Where there's a bang,
There's Yang.

Laboratory bangs can range all the way from an exploding test tube to the explosion of a hydrogen bomb. But the really Big Bangs are the bangs that occur inside the heads of theoretical physicists when they try to put together the pieces handed to them by the experimental physicists.

John Campbell, Jr., the editor of *Analog Science Fiction*, once speculated in an editorial that perhaps there was some difference in the intellectual heritage of the Western and Oriental worlds which had predisposed two Chinese physicists to question the symmetry of natural law. It is an interesting thought. I myself pointed out, in my Mathematical Games column in *Scientific American*, March 1958, that the great religious symbol of the Orient (it appears on the Korean national flag) is the circle divided asymmetrically as shown in Figure 63. The dark

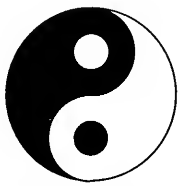


FIGURE 63. The asymmetric Yin-Yang symbol of the Orient.

and light areas are known respectively as the Yin and Yang. The Yin and Yang are symbols of all the fundamental dualities of life: good and evil, beauty and ugliness, truth and falsehood, male and female, night and day, sun and moon, heaven and earth, pleasure and pain, odd and even, left and right, positive and negative . . . the list is endless. This dualism was first symbolized in China by the odd and even digits that alternate around the perimeter of the *Lo shu*, the ancient Chinese magic square of order 3. Sometime in the tenth century the *Lo shu* was replaced by the divided circle, which soon became the dominant Yin-Yang symbol. When it was printed or drawn, black and white was used, but when painted, the Yang was made red

instead of white. The two small spots were (and still are) usually added to symbolize the fact that on each side of any duality there is always a bit of the other side. Every good act contains an element of evil, every evil act an element of good; every ugliness includes some beauty, every beauty includes some ugliness, and so on.[2] The spots remind the scientist that every "true" theory contains an element of falsehood. "Nothing is perfect," says the Philosopher in James Stephens' *The Crock of Gold*. "There are lumps in it."

EXERCISE 17: *There is a three-dimensional analog of the Yin-Yang, so familiar that almost everyone has at one time held a model of it in his hands. What is it? Is it left–right symmetrical?*

The history of science can be described as a continual, perhaps never-ending, discovery of new lumps. It was once thought that planets moved in perfect circles. Even Galileo, although he placed the sun and not the earth at the center of the solar system, could not accept Kepler's view that the planetary orbits were ellipses. Eventually it became clear that Kepler had been right: the orbits are *almost* circles but not quite. Newton's theory of gravity explained why the orbits were perfect ellipses. Then slight deviations in the Newtonian orbits turned up and were in turn explained by the correction factors of relativity theory that Einstein introduced into the Newtonian equations. "The real trouble with this world of ours," comments Gilbert Chesterton in *Orthodoxy*, "is not that it is an unreasonable world, nor even that it is a reasonable one. The commonest kind of trouble is that it is nearly reasonable, but not quite. . . . It looks just a little more mathematical and regular than it is; its exactitude is obvious, but its inexactitude is hidden; its wildness lies in wait."

To illustrate, Chesterton imagines an extraterrestrial examining a human body for the first time. He notes that the right

side exactly duplicates the left: two arms, two legs, two ears, two eyes, two nostrils, even two lobes of the brain. Probing deeper he finds a heart on the left side. He deduces that there is another heart on the right. Here of course, he encounters a spot of Yin within the Yang. "It is this silent swerving from accuracy by an inch," Chesterton continues, "that is the uncanny element in everything. It seems a sort of secret treason in the universe. . . . Everywhere in things there is this element of the quiet and incalculable."

Feynman, with no less reverence than Chesterton, says the same thing this way at the close of a lecture on symmetry in physical laws (Lecture 52 in *The Feynman Lectures on Physics,* Addison-Wesley, 1963):

"Why is nature so nearly symmetrical? No one has any idea why. The only thing we might suggest is something like this: There is a gate in Japan, a gate in Neiko, which is sometimes called by the Japanese the most beautiful gate in all Japan; it was built in a time when there was great influence from Chinese art. This gate is very elaborate, with lots of gables and beautiful carving and lots of columns and dragon heads and princes carved into the pillars, and so on. But when one looks closely he sees that in the elaborate and complex design along one of the pillars, one of the small design elements is carved upside down; otherwise the thing is completely symmetrical. If one asks why this is, the story is that it was carved upside down so that the gods will not be jealous of the perfection of man. So they purposely put the error in there, so that the gods would not be jealous and get angry with human beings.

"We might like to turn the idea around and think that the true explanation of the near symmetry of nature is this: that God made the laws only nearly symmetrical so that we should not be jealous of His perfection!"

Note that the Yin-Yang symbol is asymmetrical. It is not superposable on its mirror image. The Yin and Yang are congruent shapes, each asymmetrical, each with the same handed-

ness. By contrast the Christian symbol, the cross, is left–right symmetrical. So is the Jewish six-pointed Star of David, unless it is shown as an interlocking pair of triangles that cross alternately over and under each other. It is a pleasant thought that perhaps the familiar asymmetry of the oriental symbol, so much a part of Chinese culture, may have played a subtle, unconscious role in making it a bit easier for Lee and Yang to go against the grain of scientific orthodoxy; to propose a test which their more symmetric-minded Western colleagues had thought scarcely worth the effort.

NOTES

1. For the benefit of readers interested in recreational mathematics, I cannot resist adding that Feynman is one of the codiscoverers of hexaflexagons, those remarkable paper-folded structures that keep changing their faces when flexed. (See Chapter 1 of my *Scientific American Book of Mathematical Puzzles and Diversions.*) Although a hexaflexagon *looks* perfectly symmetrical, its inner structure possesses a handedness; that is, any given flexagon can be constructed in either a left or right-handed way.

In 1949 Feynman had suggested that perhaps the positron is an electron moving temporarily backward in time ("The Theory of Positrons," *Physical Review*, Vol. 76, 1949, pp. 749-759; reprinted in *Quantum Electrodynamics*, edited by Julius Schwinger, Dover, 1958). This prompted speculations that antiparticles are simply particles moving backward in time, and that time might be reversed (relative to our time) in galaxies of antimatter. (See "The Tiniest Time Traveler" by David Fox, *Astounding Science Fiction*, December 1952; "Speculations Concerning Precognition" by I. J. Good in his anthology of "partly baked ideas," *The Scientist Speculates*, Basic, 1962, pp. 151ff.)

It is true that if a motion picture of a spinning top is run backward, the picture will be the same as if mirror reversed, but there are strong technical reasons why time reversal cannot be invoked as an explanation of parity violation in weak interactions. Hans Reichenbach, in his book *The Direction of Time* (University of California Press, 1956, pp. 262-269), calls Feynman's positron theory "the most serious blow the concept of time has ever received in physics." Not only does it reverse the *direction* of time for parts of the world, Reichenbach points out, it

also destroys the uniform topological *order* of causal chains. Admirers of Lewis Carroll need not be reminded of the Outlandish Watch (*Sylvie and Bruno,* Chapter 23) with its "reversal-peg" that causes time to flow backward.

2. For these facts about the Yin-Yang symbol I am indebted to Schuyler Cammann's excellent article on "The Magic Square of Three in Old Chinese Philosophy and Religion," *History of Religions,* Vol. 1, No. 1, Summer 1961, pp. 37-80.

The entertaining and theoretically powerful concept of time going backward creates a variety of paradoxes.

18 Can Time Go Backward?

Martin Gardner

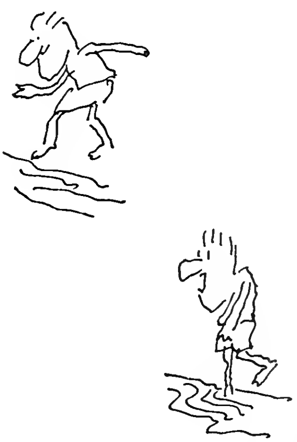
Scientific American article, published in 1967.

"...time, dark time, secret time, forever flowing like a river...."

—THOMAS WOLFE,
The Web and the Rock

Time has been described by many metaphors, but none is older or more persistent than the image of time as a river. You cannot step twice in the same river, said Heraclitus, the Greek philosopher who stressed the temporal impermanence of all things, because new waters forever flow around you. You cannot even step into it once, added his pupil Cratylus, because while you step both you and the river are changing into something different. As Ogden Nash put it in his poem "Time Marches On,"

While ladies draw their stockings on,
The ladies they were are up and gone.



RIVER IMAGE appealed to ancient Greek philosophers. You cannot step twice into the same river, said Heraclitus. Indeed, added Cratylus, you cannot do it even once.

In James Joyce's *Finnegans Wake* the great symbol of time is the river Liffey flowing through Dublin, its "hither-and-thithering waters" reaching the sea in the final lines, then returning to "riverrun," the book's first word, to begin again the endless cycle of change.

It is a powerful symbol, but also a confusing one. It is not time that flows but the world. "In what units is the rate of time's flow to be measured?" asked the Australian philosopher J. J. C. Smart. "Seconds per ——?" To say "time moves" is like saying "length extends." As Austin Dobson observed in his poem "The Paradox of Time,"

Time goes, you say? Ah no!
Alas, time stays, we go.

Moreover, whereas a fish can swim upriver against the current, we are powerless to move into the past. The changing world seems more like the magic green carpet that carried Ozma across the Deadly Desert (the void of nothingness?), unrolling only at the front, coiling up only at the back, while she journeyed from Oz to Ev, walking always in one direction on the carpet's tiny green region of "now." Why does the magic carpet never roll backward? What is the physical basis for time's strange, undeviating asymmetry?

There has been as little agreement among physicists on this matter as there has been among philosophers. Now, as the result of recent experiments, the confusion is greater than ever. Before 1964 all the fundamental laws of physics, including relativity and quantum laws, were "time-reversible." That is to say, one could substitute $-t$ for t in any basic law and the law would remain as applicable to the world as before; regardless of the sign in front of t the law described something that could occur in nature. Yet there are many events that are possible in theory but that never or almost never actually take place. It was toward those events that physicists turned their attention in the hope of finding an ultimate physical basis for distinguishing the front from the back of "time's arrow."

A star's radiation, for example, travels outward in all directions. The reverse is never observed: radiation coming from all directions and converging on a star with backward-running nuclear reactions that make it an energy sink instead of an energy source. There is nothing in the basic laws to make such a situation impossible in principle; there is only the difficulty of imagining how it could get started. One would have to assume that God or the gods, in some higher continuum, started the waves at the rim of the universe. The emergence of particles from a disintegrating radioactive nucleus and the production of ripples when a stone is dropped into a quiet lake are similar instances of one-way events. They never occur in reverse because of the enormous improbability that "boundary conditions"—conditions at the "rim" of things—would be such as to produce the required kind of converging energy. The reverse of beta decay, for instance, would require that an electron, a proton and an antineutrino be shot from the "rim" with such deadly accuracy of aim that all three particles would strike the same nucleus and create a neutron.

The steady expansion of the entire cosmos is another example. Here again there is no reason why this could not, in principle, go the other way. If the directions of all the receding galaxies were reversed, the red shift would become a blue shift, and the total picture would violate no known physical laws. All

these expanding and radiative processes, although always one-way as far as our experience goes, fail to provide a fundamental distinction between the two ends of time's arrow.

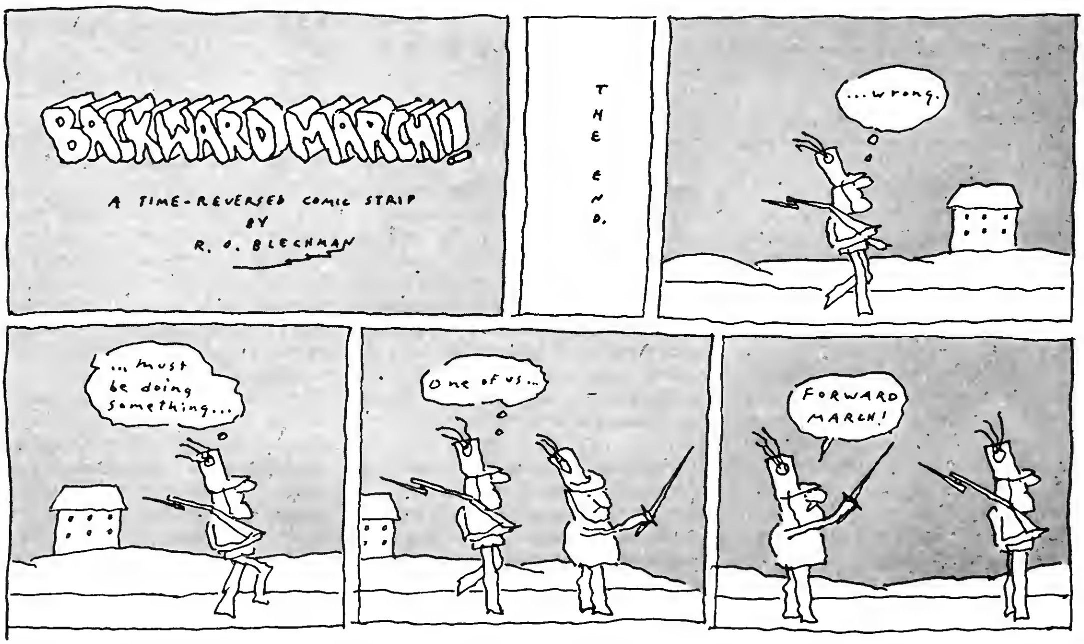
It has been suggested by many philosophers, and even by some physicists, that it is only in human consciousness, in the one-way operation of our minds, that a basis for time's arrow can be found. Their arguments have not been convincing. After all, the earth had a long history before any life existed on it, and there is every reason to believe that earthly events were just as unidirectional along the time axis then as they are now. Most physicists came finally to the conclusion that all natural events are time-reversible in principle (this became known technically as "time invariance") except for events involving the statistical behavior of large numbers of interacting objects.

Consider what happens when a cue ball breaks a triangle of 15 balls on a pool table. The balls scatter hither and thither and the 8 ball, say, drops into a side pocket. Suppose immediately after this event the motions of all the entities involved are reversed in direction while keeping the same velocities. At the spot where the 8 ball came to rest the molecules that carried off the heat and shock of impact would all converge on the same spot to create a small explosion that would start the ball back up the incline. Along the way the molecules that carried off the heat of friction would move toward the ball and boost it along its upward path. The other balls would be set in motion in a similar fashion. The 8 ball would be propelled out of the side pocket and the balls would move around the table until they finally converged to form a triangle. There would be no sound of impact because all the molecules that had been involved in the shock waves produced by the initial break of the triangle would be converging on the balls and combining with their momentum in such a way that the impact would freeze the triangle and shoot the cue ball back toward the tip of the cue. A motion picture of any individual molecule in this event would show absolutely nothing unusual. No basic mechanical law would seem to be violated. But when the billions of "hither-and-thithering" molecules involved in the total picture are considered, the probability that they would all move in the way required for the time-reversed event is so low that no one can conceive of its happening.

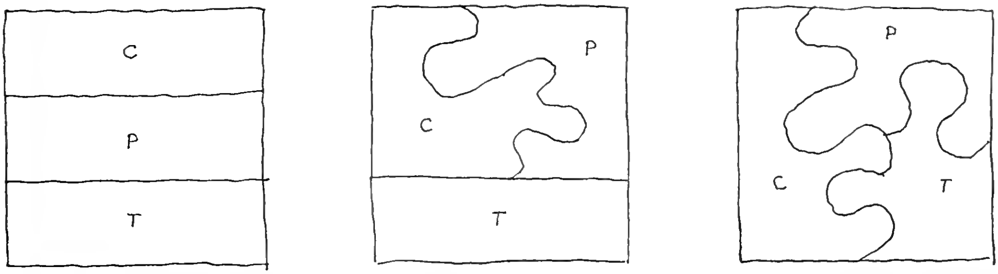
Because gravity is a one-way force, always attracting and never repelling, it might be supposed that the motions of bodies under the influence of gravity could not be time-reversed without violating basic laws. Such is not the case. Reverse the directions of the planets and they would swing around the sun in the same orbits. What about the collisions of objects drawn together by gravity—the fall of a meteorite, for example? Surely *this* event is not time-reversible. But it is! When a large meteorite strikes the earth, there is an explosion. Billions of molecules scatter hither and thither. Reverse the directions of all those molecules and their impact at one spot would provide just the right amount of energy to send the meteorite back into orbit. No basic laws would be violated, only statistical laws.

It was here, in the laws of probability, that most 19th-century physicists found an ultimate basis for time's arrow. Probability explains such irreversible processes as the mixing of coffee and cream, the breaking of a window by a stone and all the other familiar one-way-only events in which large numbers of molecules are involved. It explains the second law of thermodynamics, which says that heat always moves from hotter to cooler regions, increasing the entropy (a measure of a certain kind of disorder) of the system. It explains why shuffling randomizes a deck of ordered cards.

"Without any mystic appeal to consciousness," declared Sir Arthur Edding-



LIVING BACKWARD in a time-forward world leads to all kinds of difficulties. It is possible, however, to imagine galaxies in which time's arrow is reversed or to consider, at the level of quantum theory, that some particles may move "the wrong way" in time.



THREE SYMMETRIES, charge (*C*), parity (*P*) and time (*T*), are likened to pieces that fit into a pattern. Before 1957 they were all assumed to be symmetrical; any experiment (the pattern) involving the three could be duplicated with any one piece, any two or all three reversed (*left*). Then experiments were found that violate *P*-symmetry, suggesting that if overall (*CPT*) symmetry holds, some piece other than *P* must also be asymmetrical. *C* was found to be such a piece; an experiment remains the same if *C* and *P* are reversed together (*middle*). In 1964 experiments that violate this *CP*-symmetry were reported. It follows that *T* must be asymmetrical in these cases, since a pattern violating *CP*-symmetry can be duplicated only by reversing all three pieces simultaneously (*right*).

ton (in a lecture in which he first introduced the phrase "time's arrow"), "it is possible to find a direction of time.... Let us draw an arrow arbitrarily. If as we follow the arrow we find more and more of the random element in the state of the world, then the arrow is pointing towards the future; if the random element decreases the arrow points towards the past. That is the only distinction known to physics."

Eddington knew, of course, that there are radiative processes, such as beta decay and the light from suns, that never go the other way, but he did not consider them sufficiently fundamental to provide a basis for time's direction. Given the initial and boundary conditions necessary for starting the reverse of a radiative process, the reverse event is certain to take place. Begin with a deck of disordered cards, however, and the probability is never high that a random shuffle will separate them into spades, hearts, clubs and diamonds. Events involving shuffling processes seem to be irreversible in a stronger sense than radiative events. That is why Eddington and other physicists and philosophers argued that statistical laws provide the most fundamental way to define the direction of time.

It now appears that there is a basis for time's arrow that is even more fundamental than statistical laws. In 1964 a group of Princeton University physicists discovered that certain weak interactions of particles are apparently not time-reversible [see "Violations of Symmetry in Physics," by Eugene P. Wigner; SCIENTIFIC AMERICAN, December, 1965]. One says "apparently" because the evidence is both indirect and controversial. Although it is possible to run certain particle interactions backward to make a direct test of time symmetry, such experiments have not as yet shown any violations of time-reversibility. The Princeton tests were of an indirect kind. They imply, if certain premises are granted, that time symmetry is violated.

The most important premise is known as the *CPT* theorem. *C* stands for electric charge (plus or minus), *P* for parity (left or right mirror images) and *T* for time (forward or backward). Until a decade ago physicists believed each of these three basic symmetries held throughout nature. If you reversed the charges on the particles in a stone, so that plus charges became minus and minus charges became plus, you would still have a stone. To be sure, the stone would be made of antimatter, but there is no reason why antimatter cannot exist. An antistone on the earth would instantly explode (matter and antimatter annihilate each other when they come in contact), but physicists could imagine a galaxy of antimatter that would behave exactly like our own galaxy; indeed, it could be in all respects exactly like our own except for its *C* (charge) reversal.

The same universal symmetry was believed to hold with respect to *P* (parity). If you reversed the parity of a stone or a galaxy—that is, mirror-reflected its entire structure down to the last wave and particle—the result would be a perfectly normal stone or galaxy. Then in 1957 C. N. Yang and T. D. Lee received the Nobel prize in physics for theoretical work that led to the discovery that parity is *not* conserved [see "The Overthrow of Parity," by Philip Morrison; SCIENTIFIC AMERICAN, April, 1957]. There are events on the particle level, involving weak interactions, that cannot occur in mirror-reflected form.

It was an unexpected and disturbing blow, but physicists quickly regained their balance. Experimental evidence was found that if these asymmetrical, parity-violating events were reflected in a special kind of imaginary mirror called the *CP* mirror, symmetry was restored. If in addition to ordinary mirror reflection there is also a charge reversal, the result is something nature can "do." Perhaps there are galaxies of antimatter that are also mirror-reflected matter. In such galaxies, physicists speculated, scientists could duplicate every particle experiment that can be performed here. If we were in communication with scientists in such a *CP*-reversed galaxy, there would be no way to discover whether they were in a world like ours or in one that was *CP*-reflected. (Of course, if we went there and our spaceship exploded on arrival, we would know we had entered a region of antimatter.)

No sooner had physicists relaxed a bit with this newly restored symmetry than the Princeton physicists found some weak interactions in which *CP* symmetry appears to be violated. In different words, they found some events that, when *CP*-reversed, are (in addition to their *C* and *P* differences) not at all duplicates of each other. It is at this point that time indirectly enters the picture, because the only remaining "magic mirror" by which symmetry can be restored is the combined *CPT* mirror in which all three symmetries—charge, parity and time—are reversed. This *CPT* mirror is not just something physicists want to preserve because they love symmetry. It is built into the foundations of relativity theory in such a way that, if it turned out not to be true, relativity theory would be in serious trouble. There are therefore strong grounds for believing the *CPT* theorem holds. *On the assumption that it does,* a violation of *CP* symmetry would imply that time symmetry is also violated [*see illustration above*]. There are a few ways to preserve the *CP* mirror without combining it with *T*,

but none has met with any success. The best way is to suppose there is a "fifth force" (in addition to the four known forces: gravity, the weak-interaction force, electromagnetism and the nuclear force) that is causing the newly discovered anomalies. Experiments have cast strong doubt on the fifth-force hypothesis, however.

Early this year Paolo Franzini and his wife, working with the alternating-gradient synchrotron at the Brookhaven National Laboratory, found even stronger evidence of *CP* violations—this time in events involving electromagnetic reactions. The Franzini work was controverted, however, by a group of physicists at the European Organization for Nuclear Research (CERN) in Geneva, who announced their results in September. At the moment the cause of this discrepancy in results is not clear.

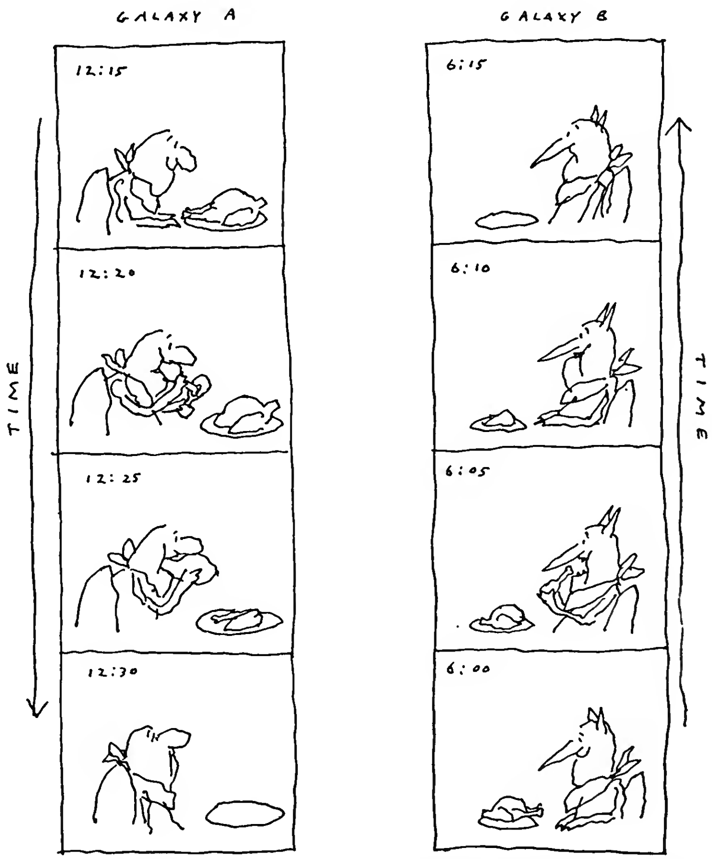
Although the evidence is still indirect and in part controversial, many physicists are now convinced that there are events at the particle level that go in only one time direction. If this holds throughout the universe, there is now a way to tell, while communicating with scientists in a distant galaxy, whether they are in a world of matter or of antimatter. We simply ask them to perform one of the *CP*-violating experiments. If their description of such a test coincides exactly with our own description of the same test when done here, we shall not explode when we visit them. It may well be that the universe contains no galaxies of antimatter. But physicists like to balance things, and if there is as much antimatter as there is matter in the universe, there may be regions of the cosmos in which all three symmetries are reversed. Events in our world that are lopsided with respect to *CPT* would all go the other way in a *CPT*-reversed galaxy. Its matter would be mirror-reflected, reversed in charge and moving backward in time.

What does it mean to say that events in a galaxy are moving backward in time? At this point no one really knows. The new experiments indicate that there is a preferred time direction for certain particle interactions. Does this arrow have any connection with other time arrows such as those that are defined by radiative processes, entropy laws and the psychological time of living organisms? Do all these arrows have to point the same way or can they vary independently in their directions?

Before the recent discoveries of the violation of *T* invariance the most popular way to give an operational meaning to "backward time" was by imagining a world in which shuffling processes went backward, from disorder to order. Ludwig Boltzmann, the 19th-century Austrian physicist who was one of the founders of statistical thermodynamics, realized that after the molecules of a gas in a closed, isolated container have reached a state of thermal equilibrium—that is, are moving in complete disorder with maximum entropy—there will always be little pockets forming here and there where entropy is momentarily decreasing. These would be balanced by other regions where entropy is increasing; the overall entropy remains relatively stable, with only minor up-and-down fluctuations.

Boltzmann imagined a cosmos of vast size, perhaps infinite in space and time, the overall entropy of which is at a maximum but which contains pockets where for the moment entropy is decreasing. (A "pocket" could include billions of galaxies and the "moment" could be billions of years.) Perhaps our fly-speck portion of the infinite sea of space-time is one in which such a fluctuation has occurred. At some time in the past, perhaps at the time of the "big bang," entropy happened to decrease; now it is increasing. In the eternal and infinite flux a bit of order happened to put in its appearance; now that order is disappearing again, and so our arrow of time runs in the familiar direction of increasing entropy. Are there other regions of space-time, Boltzmann asked, in which the arrow of entropy points the other way? If so, would it be correct to say that time in such a region was moving backward, or should one simply say that entropy was decreasing as the region continued to move forward in time?

It seems evident today that one cannot speak of backward time without meaning considerably more than just a reversal of the entropy arrow. One has



TIME IS RELATIONAL, not absolute. Observers in galaxies with opposite time directions each suppose the other to be moving backward in time. The man in *A* sees a diner in *B* eating backward; the diner in *B*, whose time is reversed, sees the man in *A* eating backward.

SHUFFLING ordinarily randomizes a pack of cards; it would be surprising to find it working the other way. Statistical laws therefore provide a way to define the direction of time.

to include all the other one-way processes with which we are familiar, such as the radiative processes and the newly discovered *CP*-violating interactions. In a world that was completely time-reversed *all* these processes would go the other way. Now, however, we must guard against an amusing verbal trap. If we imagine a cosmos running backward while we stand off somewhere in space to observe the scene, then we must be observing the cosmos moving backward in a direction opposite to our own psychological time, which still runs forward. What does it mean to say that the *entire* cosmos, including all possible observers, is running backward?

In the first book of Plato's *Statesman* a stranger explains to Socrates his theory that the world goes through vast oscillating cycles of time. At the end of each cycle time stops, reverses and then goes the other way. This is how the stranger describes one of the backward cycles:

"The life of all animals first came to a standstill, and the mortal nature ceased to be or look older, and was then reversed and grew young and delicate; the white locks of the aged darkened again, and the cheeks of the bearded man became smooth, and recovered their former bloom; the bodies of youths in their prime grew softer and smaller, continually by day and night returning and becoming assimilated to the nature of a newly born child in mind as well as body; in the succeeding stage they wasted away and wholly disappeared."

Plato's stranger is obviously caught in the trap. If things come to a standstill in time and "then" reverse, what does the word "then" mean? It has meaning only if we assume a more fundamental kind of time that continues to move forward, altogether independent of how things in the universe move. Relative to this meta-time—the time of the hypothetical observer who has slipped unnoticed into the picture—the cosmos is indeed running backward. But if there *is* no meta-time—no observer who can stand outside the entire cosmos and watch it reverse—it is hard to understand what sense can be given to the statement that the cosmos "stops" and "then" starts moving backward.

There is less difficulty—indeed, no logical difficulty at all—in imagining two portions of the universe, say two galaxies, in which time goes one way in one galaxy and the opposite way in the other. The philosopher Hans Reichenbach, in his book *The Direction of Time*, suggests that this could be the case, and that intelligent beings in each galaxy would regard their own time as "forward" and time in the other galaxy as "backward." The two galaxies would be like two mirror images: each would seem reversed to inhabitants of the other [*see illustration on preceding page*]. From this point of view time is a relational concept like up and down, left and right or big and small. It would be just as meaningless to say that the *entire* cosmos reversed its time direction as it would be to say that it turned upside down or suddenly became its own mirror image. It would be meaningless because there is no absolute or fixed time arrow outside the cosmos by which such a reversal could be measured. It is only when *part* of the cosmos is time-reversed in relation to another part that such a reversal acquires meaning.

Now, however, we come up against a significant difference between mirror reflection and time reversal. It is easy to observe a reversed world—one has only to look into a mirror. But how could an observer in one galaxy "see" another galaxy that was time-reversed? Light, instead of radiating from the other galaxy, would seem to be going toward it. Each galaxy would be totally invisible to the other. Moreover, the memories of observers in the two galaxies would be operating in opposite directions. If you somehow succeeded in communicating something to someone in a time-reversed world, he would promptly forget it because the event would instantly become part of his future rather than of his past. "It's a poor sort of memory that only works backward," said Lewis Carroll's White Queen in one looking-glass, time-reversed (*PT*-reversed!) scene. Unfortunately, outside of Carroll's dream world, memory works only one way. Norbert Wiener, speculating along such lines in his book *Cybernetics*, concluded that no communication would be possible between intelligent beings in regions with opposite time directions.

A British physicist, F. Russell Stannard, pursues similar lines of thought in an article on "Symmetry of the Time Axis" (*Nature*, August 13, 1966) and goes even further than Wiener. He concludes (and not all physicists agree with him) that no interactions of any kind would be possible between particles of matter in two worlds whose time axes pointed in opposite directions. If the universe maintains an overall symmetry with respect to time, matter of opposite time directions would "decouple" and the two worlds would become invisible to each other. The "other" world "would consist of galaxies absorbing their light rather than emitting it, living organisms growing younger, neutrons being created in triple collisions between protons, electrons and antineutrinos, and thereafter being absorbed in nuclei, etc. It would be a universe that was in a state of contraction, and its entropy would be decreasing, and yet the faustian observers ["faustian" is Stannard's term for the "other" region] would not be aware of anything strange in their environment. Being constructed of faustian matter, their subjective experience of time is reversed, so they would be equally convinced that it was they who grew older and their entropy that increased."

Instead of one universe with oscillating time directions, as in the vision of Plato's stranger, Stannard's vision bifurcates the cosmos into side-by-side regions, each unrolling its magic carpet of history simultaneously (whatever "simultaneously" can mean!) but in opposite directions. Of course, there is no reason why the cosmos has to be sym-

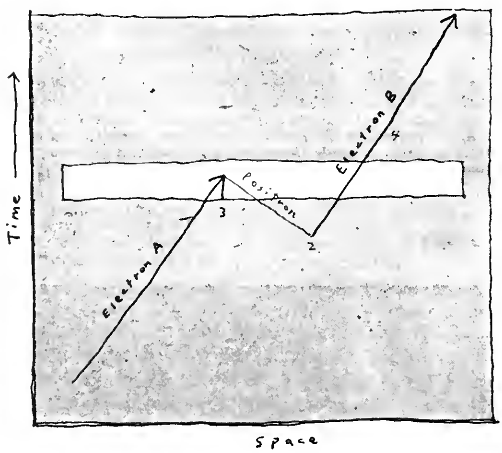
metrical in an overall way just to satisfy the physicist's aesthetic sense of balance. The universe may well be permanently lopsided in regard to all three aspects—charge, parity and time—even if there is no theoretical reason why all three could not go the other way. If a painting does not have to be symmetrical to be beautiful, why should the universe?

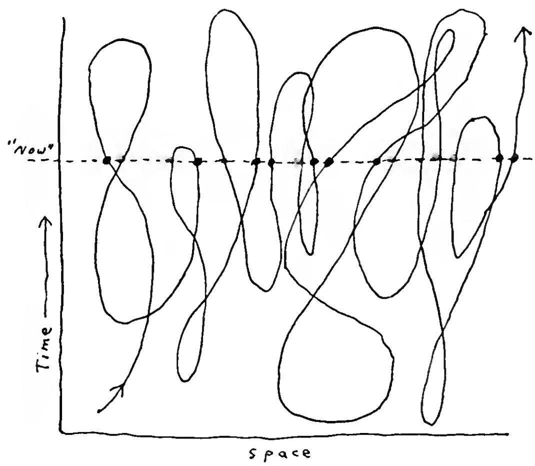
Is it possible to imagine a single individual living backward in a time-forward world? Plato's younger contemporary, the Greek historian Theopompus of Chios, wrote about a certain fruit that, when eaten, would start a person growing younger. This, of course, is not quite the same thing as a complete reversal of the person's time. There have been several science-fiction stories about individuals who grew backward in this way, including one amusing tale, "The Curious Case of Benjamin Button," by (of all people) F. Scott Fitzgerald. (It first appeared in *Colliers* in 1922 and is most accessible at the moment in *Pause to Wonder,* an anthology edited by Marjorie Fischer and Rolfe Humphries.) Benjamin is born in 1860, a 70-year-old man with white hair and a long beard. He grows backward at a normal rate, enters kindergarten at 65, goes through school and marries at 50. Thirty years later, at the age of 20, he decides to enter Harvard, and he is graduated in 1914 when he is 16. (I am giving his biological ages.) The Army promotes him to brigadier general because as a biologically older man he had been a lieutenant colonel during the Spanish-American War, but when he shows up at the Army base as a small boy he is packed off for home. He grows younger until he cannot walk or talk. "Then it was all dark," reads Fitzgerald's last sentence, "and his white crib and the dim faces that moved above him, and the warm sweet aroma of the milk, faded out altogether from his mind."

Aside from his backward growth, Mr. Button lives normally in forward-moving time. A better description of a situation in which the time arrows of a person and the world point in opposite directions is found in Carroll's novel *Sylvie and Bruno Concluded.* The German Professor hands the narrator an Outlandish Watch with a "reversal peg" that causes the outside world to run backward for four hours. There is an amusing description of a time-reversed dinner at which "an empty fork is raised to the lips: there it receives a neatly cut piece of mutton, and swiftly conveys it to the plate, where it instantly attaches itself to the mutton already there." The scene is not consistent, however. The order of the dinner-table remarks is backward, but the words occur in a forward time direction.

If we try to imagine an individual whose entire bodily and mental processes are reversed, we run into the worst kind of difficulties. For one thing, he could not pass through his previous life experiences backward, because those experiences are bound up with his external world, and since that world is still moving forward none of his past experiences can be duplicated. Would we see him go into a mad death dance, like an automaton whose motor had been reversed? Would he, from his point of view, find himself still thinking forward in a world that seemed to be going backward? If so, he would be unable to see or hear anything in that world, because all sound and light waves would be moving toward their points of origin.

We seem to encounter nothing but nonsense when we try to apply different time arrows to an individual and the world. Is it perhaps possible, on the microlevel of quantum theory, to speak sensibly about part of the universe moving the wrong way in time? It is. In 1948 Richard P. Feynman, who shared the 1965 Nobel prize in physics, developed a mathematical approach to quantum theory in which an antiparticle is regarded as a particle moving backward in time for a fraction of a microsecond. When there is pair-creation of an electron and its antiparticle the positron (a positively charged electron), the positron is extremely short-lived. It immediately collides with another electron,





FEYNMAN GRAPH, shown at the left in a simplified form devised by Banesh Hoffman of Queens College, shows how an antiparticle can be considered a particle moving backward in time. The graph is viewed through a horizontal slot in a sheet of cardboard (*color*) that is moved slowly up across the graph. Looking through the slot, one sees events as they appear in our forward-looking "now." Electron *A* moves to the right (*1*), an electron-positron pair is created (*2*), the positron and electron *A* are mutually annihilated (*3*) and electron *B* continues on to the right (*4*). From a timeless point of view (without the slotted cardboard), however, it can be seen that there is only one particle: an electron that goes forward in time, backward and then forward again. Richard P. Feynman's approach stemmed from a whimsical suggestion by John A. Wheeler of Princeton University: a single particle, tracing a "world line" through space and time (*right*), could create all the world's electrons (*black dots*) and positrons (*colored dots*).

both are annihilated and off goes a gamma ray. Three separate particles—one positron and two electrons—seem to be involved. In Feynman's theory there is only *one* particle, the electron [*see illustration on opposite page*]. What we observe as a positron is simply the electron moving momentarily back in time. Because our time, in which we observe the event, runs uniformly forward, we see the time-reversed electron as a positron. We think the positron vanishes when it hits another electron, but this is just the original electron resuming its forward time direction. The electron executes a tiny zigzag dance in space-time, hopping into the past just long enough for us to see its path in a bubble chamber and interpret it as a path made by a positron moving forward in time.

Feynman got his basic idea when he was a graduate student at Princeton, from a telephone conversation with his physics professor John A. Wheeler. In his Nobel-prize acceptance speech Feynman told the story this way:

"Feynman," said Wheeler, "I know why all electrons have the same charge and the same mass."

"Why?" asked Feynman.

"Because," said Wheeler, "they are all the *same* electron!"

Wheeler went on to explain on the telephone the stupendous vision that had come to him. In relativity theory physicists use what are called Minkowski graphs for showing the movements of objects through space-time. The path of an object on such a graph is called its "world line." Wheeler imagined one electron, weaving back and forth in space-time, tracing out a single world line. The world line would form an incredible knot, like a monstrous ball of tangled twine with billions on billions of crossings, the "string" filling the entire cosmos in one blinding, timeless instant. If we take a cross section through cosmic space-time, cutting at right angles to the time axis, we get a picture of three-space at one instant of time. This three-dimensional cross section moves forward along the time axis, and it is on this moving section of "now" that the events of the world execute their dance. On this cross section the world line of the electron, the incredible knot, would be broken up into billions on billions of dancing points, each corresponding to a spot where the electron knot was cut. If the cross section cuts the world line at a spot where the particle is moving forward in time, the spot is an electron. If it cuts the world line at a spot where the particle is moving backward in time, the spot is a positron. All the electrons and positrons in the cosmos are, in Wheeler's fantastic vision, cross sections of the knotted path of this single particle. Since they are all sections of the same world line, naturally they will all have identical masses and strengths of charge. Their positive and negative charges are no more than indications of the time direction in which the particle at that instant was weaving its way through space-time.

There is an enormous catch to all of this. The number of electrons and positrons in the universe would have to be equal. You can see this by drawing on a sheet of paper a two-dimensional analogue of Wheeler's vision. Simply trace a single line over the page to make a tangled knot [*see illustration on opposite page*]. Draw a straight line through it. The straight line represents a one-dimensional cross section at one instant in time through a two-space world (one space axis and one time axis). At points where the knot crosses the straight line, moving up in the direction of time's arrow, it produces an electron. Where it crosses the line going the opposite way it produces a positron. It is easy to see that the number of electrons and positrons must be equal or have at most a difference of one. That is why, when Wheeler had described his vision, Feynman immediately said:

"But, Professor, there aren't as many positrons as electrons."

"Well," countered Wheeler, "maybe they are hidden in the protons or something."

Wheeler was not proposing a serious theory, but the suggestion that a positron could be interpreted as an electron moving temporarily backward in time caught Feynman's fancy, and he found that the interpretation could be handled mathematically in a way that was entirely consistent with logic and all the laws of quantum theory. It became a cornerstone in his famous "space-time view" of quantum mechanics, which he completed eight years later and for which he shared his Nobel prize. The theory is equivalent to traditional views, but the zigzag dance of Feynman's particles provided a new way of handling certain calculations and greatly simplifying them. Does this mean that the positron is "really" an electron moving backward in time? No, that is only one physical interpretation of the "Feynman graphs"; other interpretations, just as valid, do not speak of time reversals. With the new experiments suggesting a mysterious interlocking of charge, parity

CP-REVERSED GALAXY (where charge is reversed and matter mirror-reflected) would be indistinguishable as such from the earth. But explorers from the earth would soon find out.

TIME-REVERSED INHABITANTS of a time-reversed world are not aware of anything strange in the environment because their own subjective experience of time is reversed.

and time direction, however, the zigzag dance of Feynman's electron, as it traces its world line through space-time, no longer seems as bizarre a physical interpretation as it once did.

At the moment no one can predict what will finally come of the new evidence that a time arrow may be built into some of the most elementary particle interactions. Physicists are taking more interest than ever before in what philosophers have said about time, thinking harder than ever before about what it means to say time has a "direction" and what connection, if any, this all has with human consciousness and will. Is history like a vast "riverrun" that can be seen by God or the gods from source to mouth, or from an infinite past to an infinite future, in one timeless and eternal glance? Is freedom of will no more than an illusion as the current of existence propels us into a future that in some unknown sense already exists? To vary the metaphor, is history a prerecorded motion picture, projected on the four-dimensional screen of our space-time for the amusement or edification of some unimaginable Audience?

Or is the future, as William James and others have so passionately argued, open and undetermined, not existing in *any* sense until it actually happens? Does the future bring genuine novelty—surprises that even the gods are unable to anticipate? Such questions go far beyond the reach of physics and probe aspects of existence that we are as little capable of comprehending as the fish in the river Liffey are of comprehending the city of Dublin.

When the first atomic bomb was nearly finished in the war-time laboratories, and before it was used, a group of physicists involved pleaded that the bomb should not be first dropped on a civilian target.

19 A Report to the Secretary of War

James Franck, Donald J. Hughes, J. I. Nickson, Eugene Rabinowitch,

Glenn T. Seaborg, Joyce C. Stearns, Leo Szilard.

June 1945.

I. Preamble

The only reason to treat nuclear power differently from all the other developments in the field of physics is the possibility of its use as a means of political pressure in peace and sudden destruction in war. All present plans for the organization of research, scientific and industrial development, and publication in the field of nucleonics are conditioned by the political and military climate in which one expects those plans to be carried out. Therefore, in making suggestions for the postwar organization of nucleonics, a discussion of political problems cannot be avoided. The scientists on this project do not presume to speak authoritatively on problems of national and international policy. However, we found ourselves, by the force of events during the last five years, in the position of a small group of citizens cognizant of a grave danger for the safety of this country as well as for the future of all the other nations, of which the rest of mankind is unaware. We therefore feel it our duty to urge that the political problems arising from the mastering of nuclear power be recognized in all their gravity, and that appropriate steps be taken for their study and the preparation of necessary decisions. We hope that the creation of the committee by the Secretary of War to deal with all aspects of nucleonics indicates that these implications have been recognized by the government. We believe that our acquaintance with the scientific elements of the situation and prolonged preoccupation with its worldwide political implications, imposes on us the obligation to offer to the committee some suggestions as to the possible solution of these grave problems.

Scientists have often before been accused of providing new weapons for the mutual destruction of nations instead of improving their well-being. It is undoubtedly true that the discovery of flying, for example, has so far brought much more misery than enjoyment and profit to humanity. However, in the past scientists could disclaim direct responsibility for the use to which mankind had put their disinterested discoveries. We feel compelled

to take a more active stand now because the success which we have achieved in the development of nuclear power is fraught with infinitely greater dangers than were all the inventions of the past. All of us familiar with the present state of nucleonics live with the vision before our eyes of sudden destruction visited on our own country, of a Pearl Harbor disaster repeated in thousand-fold magnification in every one of our major cities.

In the past, science has often been able to also provide new methods of protection against new weapons of aggression it made possible, but it cannot promise such efficient protection against the destructive use of nuclear power. This protection can come only from the political organization of the world. Among all the arguments calling for an efficient international organization for peace, the existence of nuclear weapons is the most compelling one. In the absence of an international authority which would make all resort to force in international conflicts impossible, nations could still be diverted from a path which must lead to total mutual destruction by a specific international agreement barring a nuclear armaments race.

II. Prospects of Armaments Race

It could be suggested that the danger of destruction by nuclear weapons can be avoided—at least as far as this country is concerned—either by keeping our discoveries secret for an indefinite time, or else by developing our nuclear armaments at such a pace that no other nation would think of attacking us from fear of overwhelming retaliation.

The answer to the first suggestion is that although we undoubtedly are at present ahead of the rest of the world in this field, the fundamental facts of nuclear power are a subject of common knowledge. British scientists know as much as we do about the basic wartime progress of nucleonics—if not of the specific processes used in our engineering developments—and the role which French nuclear physicists have played in the pre-war development of this field, plus their occasional contact with our projects, will enable them to catch up rapidly, at least as far as basic scientific discoveries are concerned. German scientists, in whose discoveries the whole development of this field originated, apparently did not develop it during the war to the same extent to which this has been done in America, but to the last day of the European war we were living in constant apprehension as to their possible achievements. The certainty that German scientists were working on this weapon and that their government would certainly have no scruples against using it when available was the main motivation of the initiative which American scientists took in urging the development of nuclear power for military purposes on a large scale in this country. In Russia, too, the basic facts and implications of nuclear power were well understood in 1940, and the experience of Russian scientists in nuclear research is entirely sufficient to enable them to retrace our steps

within a few years, even if we should make every attempt to conceal them. Even if we can retain our leadership in basic knowledge of nucleonics for a certain time by maintaining secrecy as to all results achieved on this and associated projects, it would be foolish to hope that this can protect us for more than a few years.

It may be asked whether we cannot prevent the development of military nucleonics in other countries by a monopoly on the raw materials of nuclear power. The answer is that even though the largest now known deposits of uranium ores are under the control of powers which belong to the "western" group (Canada, Belgium and British India), the old deposits in Czechoslovakia are outside this sphere. Russia is known to be mining radium on its own territory, and even if we do not know the size of the deposits discovered so far in the USSR, the probability that no large reserves of uranium will be found in a country which covers one-fifth of the land area of the earth (and whose sphere of influence takes in additional territory), is too small to serve as a basis for security. *Thus, we cannot hope to avoid a nuclear armament race either by keeping secret from the competing nations the basic scientific facts of nuclear power or by cornering the raw materials required for such a race.*

We now consider the second of the two suggestions made at the beginning of this section, and ask whether we could not feel ourselves safe in a race of nuclear armaments by virtue of our greater industrial potential, including greater diffusion of scientific and technical knowledge, greater volume and efficiency of our skilled labor corps, and greater experience of our management—all the factors whose importance has been so strikingly demonstrated in the conversion of this country into an arsenal of the allied nations in the present war. The answer is that all that these advantages can give us is the accumulation of a larger number of bigger and better atomic bombs.

However, such a quantitative advantage in reserves of bottled destructive power will not make us safe from sudden attack. Just because a potential enemy will be afraid of being "outnumbered and outgunned," the temptation for him may be overwhelming to attempt a sudden unprovoked blow—particularly if he should suspect us of harboring aggressive intentions against his security or his sphere of influence. In no other type of warfare does the advantage lie so heavily with the aggressor. He can place his "infernal machines" in advance in all our major cities and explode them simultaneously, thus destroying a major part of our industry and a large part of our population aggregated in densely populated metropolitan districts. Our possibilities of retaliation—even if retaliation should be considered adequate compensation for the loss of millions of lives and destruction of our largest cities—will be greatly handicapped because we must rely on aerial transportation of the bombs, and also because we may have to deal with an enemy whose industry and population are dispersed over a large territory.

In fact, if the race for nuclear armaments is allowed to develop, the only apparent way in which our country can be protected from the paralyzing effects of a sudden attack is by dispersal of those industries which are essential for our war efforts and dispersal of the populations of our major metropolitan cities. As long as nuclear bombs remain scarce (i.e., as long as uranium remains the only basic material for their fabrication), efficient dispersal of our industry and the scattering of our metropolitan population will considerably decrease the temptation to attack us by nuclear weapons.

At present, it may be that atomic bombs can be detonated with an effect equal to that of 20,000 tons of TNT. One of these bombs could then destroy something like three square miles of an urban area. Atomic bombs containing a larger quantity of active material but still weighing less than one ton may be expected to be available within ten years which could destroy over ten square miles of a city. A nation able to assign ten tons of atomic explosives for a sneak attack on this country can then hope to achieve the destruction of all industry and most of the population in an area from 500 square miles upwards. If no choice of targets, with a total area of 500 square miles of American territory, contains a large enough fraction of the nation's industry and population to make their destruction a crippling blow to the nation's war potential and its ability to defend itself, then the attack will not pay and may not be undertaken. At present, one could easily select in this country a hundred areas of five square miles each whose simultaneous destruction would be a staggering blow to the nation. Since the area of the United States is about three million square miles, it should be possible to scatter its industrial and human resources in such a way as to leave no 500 square miles important enough to serve as a target for nuclear attack.

We are fully aware of the staggering difficulties involved in such a radical change in the social and economic structure of our nation. We felt, however, that the dilemma had to be stated, to show what kind of alternative methods of protection will have to be considered if no successful international agreement is reached. It must be pointed out that in this field we are in a less favorable position than nations which are either now more diffusely populated and whose industries are more scattered, or whose governments have unlimited power over the movement of population and the location of industrial plants.

If no efficient international agreement is achieved, the race for nuclear armaments will be on in earnest not later than the morning after our first demonstration of the existence of nuclear weapons. After this, it might take other nations three or four years to overcome our present head start, and eight or ten years to draw even with us if we continue to do intensive work in this field. This might be all the time we would have to bring about the relocation of our population and industry. Obviously, no time should be lost in inaugurating a study of this problem by experts.

III. Prospects of Agreement

The consequences of nuclear warfare, and the type of measures which would have to be taken to protect a country from total destruction by nuclear bombing must be as abhorrent to other nations as to the United States. England, France, and the smaller nations of the European continent, with their congeries of people and industries, would be in a particularly desperate situation in the face of such a threat. Russia and China are the only great nations at present which could survive a nuclear attack. However, even though these countries may value human life less than the peoples of Western Europe and America, and even though Russia, in particular, has an immense space over which its vital industries could be dispersed and a government which can order this dispersion the day it is convinced that such a measure is necessary—there is no doubt that Russia, too, will shudder at the possibility of a sudden disintegration of Moscow and Leningrad, almost miraculously preserved in the present war, and of its new industrial cities in the Urals and Siberia. Therefore, only lack of mutual *trust* and not lack of *desire* for agreement can stand in the path of an efficient agreement for the prevention of nuclear warfare. The achievement of such an agreement will thus essentially depend on the integrity of intentions and readiness to sacrifice the necessary fraction of one's own sovereignty by all the parties to the agreement.

One possible way to introduce nuclear weapons to one world—which may particularly appeal to those who consider nuclear bombs primarily as a secret weapon developed to help win the present war—is to use them without warning on appropriately selected objects in Japan.

Although important tactical results undoubtedly can be achieved by a sudden introduction of nuclear weapons, we nevertheless think that the question of the use of the very first available atomic bombs in the Japanese war should be weighed very carefully, not only by military authorities but by the highest political leadership of this country.

Russia, and even allied countries which bear less mistrust of our ways and intentions, as well as neutral countries may be deeply shocked by this step. It may be very difficult to persuade the world that a nation which was capable of secretly preparing and suddenly releasing a new weapon as indiscriminate as the rocket bomb and a thousand times more destructive is to be trusted in its proclaimed desire of having such weapons abolished by international agreement. We have large accumulations of poison gas but do not use them, and recent polls have shown that public opinion in this country would disapprove of such a use even if it would accelerate the winning of the Far Eastern war. It is true that some irrational element in mass psychology makes gas poisoning more revolting than blasting by explosives, even though gas warfare is in no way more "inhuman" than the

war of bombs and bullets. Nevertheless, it is not at all certain that American public opinion, if it could be enlightened as to the effect of atomic explosives, would approve of our own country being the first to introduce such an indiscriminate method of wholesale destruction of civilian life.

Thus, from the "optimistic" point of view—looking forward to an international agreement on the prevention of nuclear warfare—the military advantages and the saving of American lives achieved by the sudden use of atomic bombs against Japan may be outweighed by the ensuing loss of confidence and by a wave of horror and repulsion sweeping over the rest of the world and perhaps even dividing public opinion at home.

From this point of view, a demonstration of the new weapon might best be made, before the eyes of representatives of all the United Nations, on the desert or a barren island. The best possible atmosphere for the achievement of an international agreement could be achieved if America could say to the world, "You see what sort of a weapon we had but did not use. We are ready to renounce its use in the future if other nations join us in this renunciation and agree to the establishment of an efficient international control."

After such a demonstration the weapon might perhaps be used against Japan if the sanction of the United Nations (and of public opinion at home) were obtained, perhaps after a preliminary ultimatum to Japan to surrender or at least to evacuate certain regions as an alternative to their total destruction. This may sound fantastic, but in nuclear weapons we have something entirely new in order of magnitude of destructive power, and if we want to capitalize fully on the advantage their possession gives us, we must use new and imaginative methods.

It must be stressed that if one takes the pessimistic point of view and discounts the possibility of an effective international control over nuclear weapons at the present time, then the advisability of an early use of nuclear bombs against Japan becomes even more doubtful—quite independent of any humanitarian considerations. If an international agreement is not concluded immediately after the first demonstration, this will mean a flying start toward an unlimited armaments race. If this race is inevitable, we have every reason to delay its beginning as long as possible in order to increase our head start still further.

The benefit to the nation and the saving of American lives in the future achieved by renouncing an early demonstration of nuclear bombs and letting the other nations come into the race only reluctantly, on the basis of guesswork and without definite knowledge that the "thing does work," may far outweigh the advantages to be gained by the immediate use of the first and comparatively inefficient bombs in the war against Japan. On the other hand, it may be argued that without an early demonstration it may prove difficult to obtain adequate support for further intensive development of nucleonics in this country and that thus the time gained by the postponement of an open armaments race will not be properly used.

Furthermore one may suggest that other nations are now or will soon be not entirely unaware of our present achievements, and that consequently the postponement of a demonstration may serve no useful purpose as far as the avoidance of an armaments race is concerned and may only create additional mistrust, thus worsening rather them improving the chances of an ultimate accord on the international control of nuclear explosives.

Thus, if the prospects of an agreement will be considered poor in the immediate future, the pros and cons of an early revelation of our possession of nuclear weapons to the world—not only by their actual use against Japan but also by a prearranged demonstration—must be carefully weighed by the supreme political and military leadership of the country, and the decision should not be left to the considerations of military tactics alone.

One may point out that scientists themselves have initiated the development of this "secret weapon" and it is therefore strange that they should be reluctant to try it out on the enemy as soon as it is available. The answer to this question was given above—the compelling reason for creating this weapon with such speed was our fear that Germany had the technical skill necessary to develop such a weapon, and that the German government had no moral restraints regarding its use.

Another argument which could be quoted in favor of using atomic bombs as soon as they are available is that so much taxpayers' money has been invested in these projects that the Congress and the American public will demand a return for their money. The attitude of American public opinion, mentioned earlier in the matter of the use of poison gas against Japan, shows that one can expect the American public to understand that it is sometimes desirable to keep a weapon in readiness for use only in extreme emergency; and as soon as the potentialities of nuclear weapons are revealed to the American people, one can be sure that they will support all attempts to make the use of such weapons impossible.

Once this is achieved, the large installations and the accumulation of explosive material at present earmarked for potential military use will become available for important peacetime developments, including power production, large engineering undertakings, and mass production of radioactive materials. In this way, the money spent on wartime development of nucleonics may become a boon for the peacetime development of national economy.

IV. Methods of International Control

We now consider the question of how an effective international control of nuclear armaments can be achieved. This is a difficult problem, but we think it soluble. It requires study by statesmen and international lawyers, and we can offer only some preliminary suggestions for such a study.

Given mutual trust and willingness on all sides to give up a certain part

of their sovereign rights by admitting international control of certain phases of national economy, the control could be exercised (alternatively or simultaneously) on two different levels.

The first and perhaps simplest way is to ration the raw materials—primarily the uranium ores. Production of nuclear explosives begins with the processing of large quantities of uranium in large isotope separation plants or huge production piles. The amounts of ore taken out of the ground at different locations could be controlled by resident agents of the international control board, and each nation could be allotted only an amount which would make large scale separation of fissionable isotopes impossible.

Such a limitation would have the drawback of making impossible also the development of nuclear power for peacetime purposes. However, it need not prevent the production of radioactive elements on a scale sufficient to revolutionize the industrial, scientific, and technical use of these materials, and would thus not eliminate the main benefits which nucleonics promises to bring to mankind.

An agreement on a higher level, involving more mutual trust and understanding, would be to allow unlimited production but keep exact bookkeeping on the fate of each pound of uranium mined. If in this way, check is kept on the conversion of uranium and thorium ore into pure fissionable materials, the question arises as to how to prevent accumulation of large quantities of such materials in the hands of one or several nations. Accumulations of this kind could be rapidly converted into atomic bombs if a nation should break away from international control. It has been suggested that a compulsory denaturation of pure fissionable isotopes may be agreed upon—by diluting them after production with suitable isotopes to make them useless for military purposes, while retaining their usefulness for power engines.

One thing is clear: any international agreement on prevention of nuclear armaments must be backed by actual and efficient controls. No paper agreement can be sufficient since neither this or any other nation can stake its whole existence on trust in other nations' signatures. Every attempt to impede the international control agencies would have to be considered equivalent to denunciation of the agreement.

It hardly needs stressing that we as scientists believe that any systems of control envisaged should leave as much freedom for the peacetime development of nucleonics as is consistent with the safety of the world.

V. Summary

The development of nuclear power not only constitutes an important addition to the technological and military power of the United States, but also creates grave political and economic problems for the future of this country.

Nuclear bombs cannot possibly remain a "secret weapon" at the exclusive disposal of this country for more than a few years. The scientific facts on which their construction is based are well known to scientists of other countries. Unless an effective international control of nuclear explosives is instituted, a race for nuclear armaments is certain to ensue following the first revelation of our possession of nuclear weapons to the world. Within ten years other countries may have nuclear bombs, each of which, weighing less than a ton, could destroy an urban area of more than ten square miles. In the war to which such an armaments race is likely to lead, the United States, with its agglomeration of population and industry in comparatively few metropolitan districts, will be at a disadvantage compared to nations whose population and industry are scattered over large areas.

We believe that these considerations make the use of nuclear bombs for an early unannounced attack against Japan inadvisable. If the United States were to be the first to release this new means of indiscriminate destruction upon mankind, she would sacrifice public support throughout the world, precipitate the race for armaments, and prejudice the possibility of reaching an international agreement on the future control of such weapons.

Much more favorable conditions for the eventual achievement of such an agreement could be created if nuclear bombs were first revealed to the world by a demonstration in an appropriately selected uninhabited area.

In case chances for the establishment of an effective international control of nuclear weapons should have to be considered slight at the present time, then not only the use of these weapons against Japan but even their early demonstration may be contrary to the interests of this country. A postponement of such a demonstration will have in this case the advantage of delaying the beginning of the nuclear armaments race as long as possible.

If the government should decide in favor of an early demonstration of nuclear weapons, it will then have the possibility of taking into account the public opinion of this country and of the other nations before deciding whether these weapons should be used against Japan. In this way, other nations may assume a share of responsibility for such a fateful decision.

Because of the central position of science in our civilization, physicists should be deeply concerned with the involvement of science in worldwide cultural and political affairs.

20 The Privilege of Being a Physicist

Victor F. Weisskopf

Article in *Physics Today*, 1969.

THERE ARE CERTAIN obvious privileges that a physicist enjoys in our society. He is reasonably paid; he is given instruments, laboratories, complicated and expensive machines, and he is asked not to make money with these tools, like most other people, but to spend money. Furthermore he is supposed to do what he himself finds most interesting, and he accounts for what he spends to the money givers in the form of progress reports and scientific papers that are much too specialized to be understood or evaluated by those who give the money—the federal authorities and, in the last analysis, the taxpayer. Still, we believe that the pursuit of science by the physicist is important and should be supported by the public. In order to prove this point, we will have to look deeper into the question of the relevance of science to society as a whole. We will not restrict ourselves to physics only; we will consider the relevance of all the natural sciences, but we will focus our attention on basic sciences, that is to those scientific activities that are performed without a clear practical application in mind.

The question of the relevance of scientific research is particularly important today, when society is confronted with a number of immediate urgent problems. The world is facing threats of nuclear war, the dangers of overpopulation, of a world famine, mounting social and racial conflicts, and the destruction of our natural environment by the byproducts of ever-increasing applications of technology. Can we afford to continue scientific research in view of these problems?

I will try to answer this question affirmatively. It will be the trend of my comments to emphasize the diversity in the relations between science and society; there are many sides and many aspects, each of different character, but of equal importance. We can divide these aspects into two distinct groups. On the one hand, science is important in shaping our *physical* environment; on the other, in shaping our *mental* environment. The first refers to the influence of science on technology, the second to the influence on philosophy, on our way of thinking.

Technology

The importance of science as a basis of technology is commonplace. Obviously, knowledge as to how nature works can be used to obtain power over nature. Knowledge acquired by basic science yielded a vast technical return. There is not a single industry

today that does not make use of the results of atomic physics or of modern chemistry. The vastness of the return is illustrated by the fact that the total cost of all basic research, from Archimedes to the present, is less than the value of ten days of the world's present industrial production.

We are very much aware today of some of the detrimental effects of the ever increasing pace of technological development. These effects begin to encroach upon us in environmental pollution of all kinds, in mounting social tensions caused by the stresses and dislocations of a fast changing way of life and, last but not least, in the use of modern technology to invent and construct more and more powerful weapons of destruction.

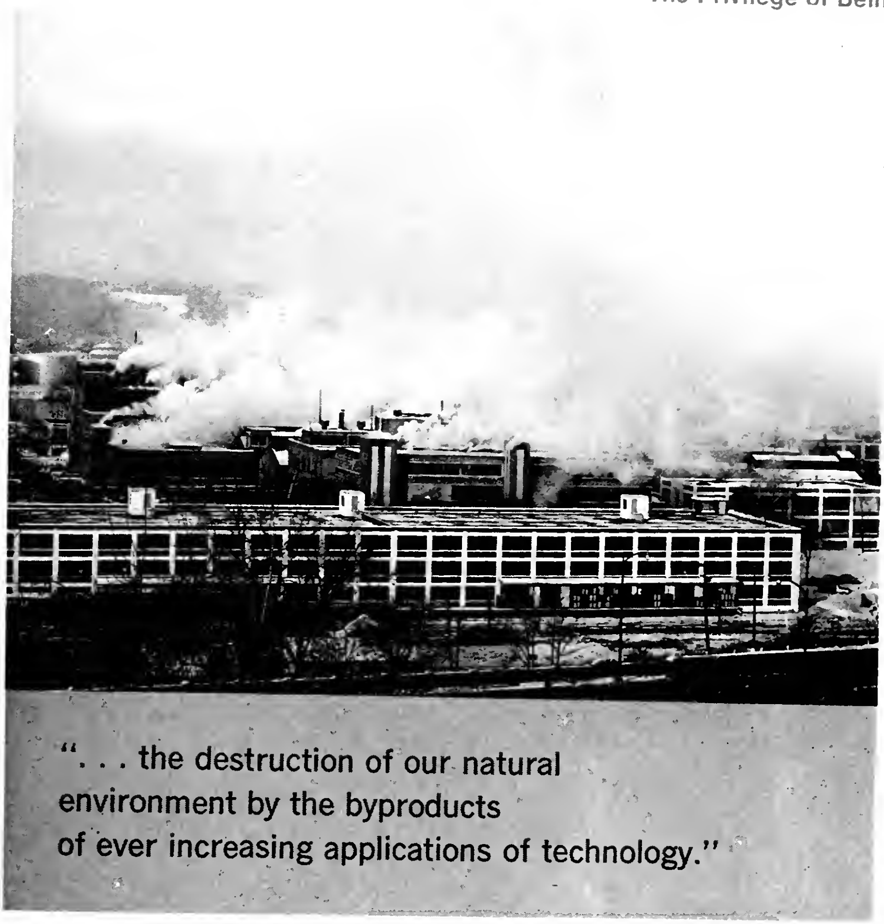
In many instances, scientific knowledge has been and should continue to be applied to counteract these effects. Certainly, physics and chemistry are useful to combat many forms of pollution and to improve public transportation. Biological research could and must be used to find more effective means of birth control and new methods to increase our food resources. It has been pointed out many times that our exploitation of the sea for food gathering is still in the hunting stage; we have not yet reached the neolithic age of agriculture and animal breeding in relation to the oceans.

Many of the problems that technology has created cannot be solved by natural science. They are social and political problems, dealing with the behavior of man in complicated and rapidly evolving situations. In particular, the questions arise: "What technical possibilities should or should not be realized? How far should they be developed?" A systematic investigation of the positive and negative social effects of technical innovations is necessary. But it is only partly a problem for natural sciences; to a greater extent, it is a problem of human behavior and human reaction. I am thinking here of the supersonic transport, of space travel, of the effects of the steadily increasing automobile traffic and again, last but not least, of the effects of the development of weapons of mass destruction.

Physical environment

What role does basic science have in shaping our physical environment? It is often said that modern basic physical science is so advanced that its problems have little to do with our terrestrial environment. It is interested in nuclear and subnuclear phenomena and in the physics of extreme

After taking his PhD at Göttingen in 1931, Victor F. Weisskopf worked at Berlin, Copenhagen, Zürich, Rochester and Los Alamos. He joined Massachusetts Institute of Technology in 1945 and has been there ever since, apart from a five-year leave of absence (1961–65) when he was director-general of CERN in Geneva. In 1956 he received the Max Planck medal for his work in theoretical physics, and he is currently head of the physics department at MIT and chairman of the high-energy physics advisory panel to AEC's research division.



temperatures. These are objectives relating to cosmic environments, far away from our own lives. Hence, the problems are not relevant for society; they are too far removed; they are studied for pure curiosity only. We will return later to the value of pure curiosity.

Let us first discuss how human environment is defined. Ten thousand years ago, metals were not part of human environment; pure metals are found only very rarely on earth. When man started to produce them, they were first considered as most esoteric and irrelevant materials and were used only for decoration purposes during thousands of years. Now they are an essential part of our environment. Electricity went through the same development, only much faster. It is observed naturally only in a few freak phenomena, such as lightning or friction electricity, but today it is an essential feature of our lives.

This shift from periphery to center was most dramatically exhibited in nuclear physics. Nuclear phenomena are certainly far removed from our terrestrial world. Their place in nature is found rather in the center of stars or of exploding supernovae, apart from a few naturally radioactive materials which are the last embers of the cosmi explosion in which terrestrial matter was formed. This is why Ernest Rutherford remarked in 1927, "Anyone who expects a source of power from

transformations of atoms is talking moonshine." It is indeed a remarkable feat to recreate cosmic phenomena on earth as we do with our accelerators and reactors, a fact often overlooked by the layman, who is more impressed by rocket trips to the moon. That these cosmic processes can be used for destructive as for constructive purposes is more proof of their relevance in our environment.

Even phenomena as far removed from daily life as those discovered by high-energy physicists may some day be of technical significance. Mesons and hyperons are odd and rare particles today, but they have interactions with ordinary matter. Who knows what these interactions may be used for at the end of this century? Scientific research not only investigates our natural environment, it also creates new artificial environments, which play an ever-increasing role in our lives.

Mental environment

The second and most important aspect of the relevance of science is its influence on our thinking, its shaping of our mental environment. One frequently hears the following views as to the effect of science on our thought: "Science is materialistic, it reduces all human experience to material processes, it undermines moral, ethical and aesthetic values because it does not recognize them, as they cannot be expressed in numbers. The world of nature is dehumanized, relativized; there are no absolutes any more; nature is regarded as an abstract formula; things and objects are nothing but vibrations of an abstract mathematical concept . . ." (Science is accused at the same time of being materialistic and of negating matter.)

Actually science gives us a unified, rational view of nature; it is an eminently successful search for fundamental laws with universal validity; it is an unfolding of the basic processes and principles from which all natural happenings are derived, a search for the absolutes, for the invariants that govern natural processes. It finds law and order—if I am permitted to use that expression in this context—in a seemingly arbitrary flow of events. There is a great fascination in recognizing the essential features of nature's structure, and a great intellectual beauty in the compact and all-embracing formulation of a physical law. Science is a search for meaning in what is going on in the natural world, in the history of the universe, its beginnings and its possible future.

Public awareness

These growing insights into the workings of nature are not only open to the scientific expert, they are also relevant to the nonscientist. Science did create an awareness among people of all ways of life that universal natural laws exist, that the universe is not run by magic, that we are not at the mercy of a capricious universe, that the structure of matter is largely known, that life has developed slowly from inorganic matter by evolution in a period of several thousand million years, that this evolution is a unique experiment of nature here on earth, which leaves us humans with a responsibility not to spoil it. Certainly the ideas of cosmology, biology, paleontology and anthropology changed the ideas of the average man in respect to future and past. The concept of an unchanging world or a world subject to arbitrary cycles of changes is replaced by a world that continuously develops from

more primitive to more sophisticated organization.

Although there is a general awareness of the public in all these aspects of science, much more could be and must be done to bring the fundamental ideas nearer to the intelligent layman. Popularization of science should be one of the prime duties of a scientist and not a secondary one as it is now. A much closer collaboration of scientists and science writers is necessary. Seminars, summer schools, direct participation in research should be the rule for science writers, in order to obtain a free and informal contact of minds between science reporters and scientists on an equal level, instead of an undirected flow of undigested information.

PHOTO BY ROSEMARY GAFFNEY

"There is not a single industry today that does not make use of the results of atomic physics or of modern chemistry."

Education

Science also shapes our thinking by means of its role in education. The study of open scientific frontiers where unsolved fundamental problems are faced is, and should be, a part of higher education. It fosters a spirit of inquiry; it lets the student participate in the joy of a new insight, in the inspiration of new understanding. The questioning of routine methods, the search for new and untried ways to accomplish things, are important elements to bring to any problem, be it one of science or otherwise. Basic research must be an essential part of higher education. In elementary education, too, science should and does play an increasing role. Intelligent play with simple, natural phenomena, the joys of discovery of unexpected experiences, are much better ways of learning to think than any teaching by rote.

A universal language . . .

The international aspect of science should not be forgotten as an important part of its influence on our mental environment. Science is a truly human concern; its concepts and its language are the same for all human beings. It transcends any cultural and political boundaries. Scientists understand each other immediately when they talk about their scientific problems, and it is thus easier for them to speak to each other on political or cultural questions and problems about which they may have divergent opinions. The scientific community serves as a bridge across boundaries, as a spearhead of international understanding.

As an example, we quote the Pugwash meetings, where scientists from the East and West met and tried to clarify some of the divergences regarding political questions that are connected with science and technology. These meetings have contributed to a few steps that were taken towards peace, such as the stopping of bomb tests, and they prepared the ground for more rational discussions of arms control. Another example is the western European laboratory for nuclear research in Geneva—CERN—in which 12 nations collaborate successfully in running a most active center for fundamental research. They have created a working model of the United States of Europe as far as high-energy physics is concerned. It is significant that this laboratory has very close ties with the laboratories in the east European countries; CERN is also equipping and participating in experiments carried out together with Russian physicists at the new giant accelerator in Serpukhov near Moscow.

. . . occasionally inadequate

The influence of science on our thinking is not always favorable. There are dangers stemming from an uncritical application of a method of thinking, so incredibly successful in natural science, to problems for which this method is inadequate. The great success of the quantitative approach in the exploration of nature may well lead to an overstressing of this method to other problems. A remark by M. Fierz in Zurich is incisive: He said that science illuminates part of our experience with such glaring intensity that the rest remains in even deeper darkness. The part in darkness has to do with the irrational and the affective in human behavior, the realm of the emotional, the instinctive world. There are aspects of human experience to which the methods of natural science

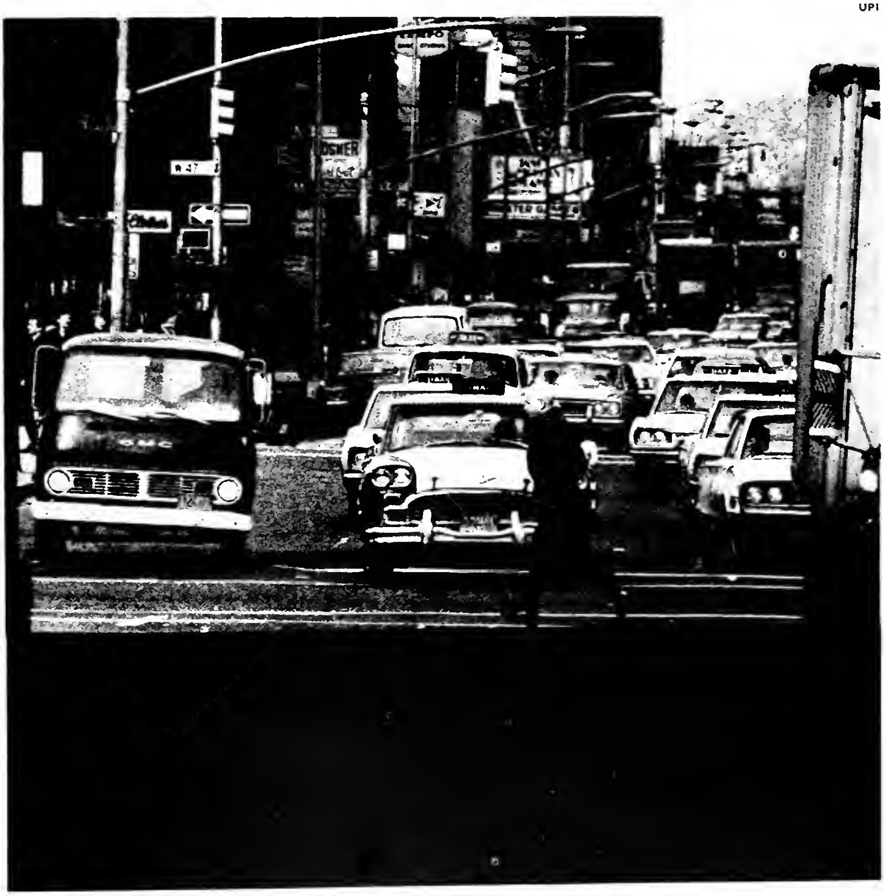
are not applicable. Seen within the framework of that science, these phenomena exhibit a degree of instability, a multidimensionality for which our present scientific thinking is inadequate and, if applied, may become dangerously misleading.

Deep involvement, deep concern

The foregoing should have served to illustrate the multilateral character of science in its relation to society. The numerous and widely differing aspects of relevance emphasize the central position of science in our civilization. Here we find a real privilege of being a scientist. He is in the midst of things; his work is deeply involved in what happens in our time. This is why it is also his privilege to be deeply concerned with the involvement of science in the events of the day.

In most instances he cannot avoid being drawn in one form or another into the decision-making process regarding the applications of science, be it on the military or on the industrial scene. He may have to help, to advise or to protest, whatever the case may be. There are different ways in which the scientist will get involved in public affairs; he may address himself to the public when he feels that science has been misused or falsely ap-

UPI



plied; he may work with his government on the manner of application of scientific results to military or social problems.

In all these activities he will be involved with controversies that are not purely scientific but political. In facing such problems and dilemmas, he will miss the sense of agreement that prevails in scientific discussions, where there is an unspoken understanding of the criteria of truth and falsehood even in the most heated controversies. Mistakes in science can easily be corrected; mistakes in public life are much harder to undo because of the highly unstable and nonlinear character of human relations.

How much emphasis?

Let us return to the different aspects of relevance in science. In times past, the emphasis has often shifted from one aspect to the other. For example at the end of the last century there was a strong overemphasis on the practical application of science in the US. Henry A. Rowland, who was the first president of the American Physical Society, fought very hard against the underemphasis of science as is seen in the following quotation from his address to the American Association for the Advancement of Science in 1883:

> "American science is a thing of the future, and not of the present or past; and the proper course of one in my position is to consider what must be done to create a science of physics in this country, rather than to call telegraphs, electric lights, and such conveniences by the name of science. I do not wish to underrate the value of all these things; the progress of the world depends on them, and he is to be honored who cultivates them successfully. So also the cook, who invents a new and palatable dish for the table, benefits the world to a certain degree; yet we do not signify him by the name of a chemist. And yet it is not an uncommon thing, especially in American newspapers, to have the applications of science confounded with pure science; and some obscure character who steals the ideas of some great mind of the past, and enriches himself by the application of the same to domestic uses, is often lauded above the great originator of the idea, who might have worked out hundreds of such applications, had his mind possessed the necessary element of vulgarity."

Rowland did succeed in his aim, although posthumously. He should have lived to see the US as the leading country in basic science for the last four decades. His statement—notwithstanding its forceful prose—appears to us today inordinately strong in its contempt of the applied physicists. The great success of this country in basic science derives to a large extent from the close coöperation of basic science with applied science. This close relation—often within the same person—provided tools of high quality, without which many fundamental discoveries could not have been made. There was a healthy equilibrium between basic and applied science during the last decades and thus also between the different aspects of the relevance of science.

Lately, however, the emphasis is changing again. There is a trend among the public, and also among scientists, away from basic science towards the application of science to immediate problems and technological shortcomings, revealed by the crisis of the day. Basic science is considered

"Intelligent play with simple, natural phenomena, the joys of discovery of unexpected experiences, are much better ways of learning to think than any teaching by rote."

to be a luxury by the public; many students and researchers feel restless in pursuing science for its own sake.

Perspective

The feeling that something should be done about the pressing social needs is very healthy. "We are in the midst of things," and scientists must face their responsibilities by using their knowledge and influence to rectify the detrimental effects of the misuse of science and technology. But we must not lose our perspective in respect to other aspects of science. We have built this great edifice of knowledge; let us not neglect it during a time of crisis. The scientist who today devotes his time to the solution of our social and environmental problems does an important job. But so does his colleague who goes on in the pur-

suit of basic science. We need basic science not only for the solution of practical problems but also to keep alive the spirit of this great human endeavor. If our students are no longer attracted by the sheer interest and excitement of the subject, we were delinquent in our duty as teachers. We must make this world into a decent and livable world, but we also must create values and ideas for people to live and to strive for. Arts and sciences must not be neglected in times of crisis; on the contrary, more weight should be given to the creation of aims and values. It is a great human value to study the world in which we live and to broaden the horizon of knowledge.

These are the privileges of being a scientist: We are participating in a most exhilarating enterprise right at the center of our culture. What we do is essential in shaping our physical and mental environment. We, therefore, carry a responsibility to take part in the improvement of the human lot and to be concerned about the consequences of our ideas and their applications. This burden makes our lives difficult and complicated and puts us in the midst of social and political life and strife.

But there are compensations. We are all working for a common and well defined aim: to get more insight into the workings of nature. It is a constructive endeavor, where we build upon the achievements of the past; we improve but never destroy the ideas of our predecessors.

This is why we are perhaps less prone to the feeling of aimlessness and instability that is observed in so many segments of our society. The growing insight into nature is not only a source of satisfaction for us, it also gives our lives a deeper meaning. We are a "happy breed of men" in a world of uncertainty and bewilderment.

* * *

This article was adapted from an address given at the joint annual meeting of the American Physical Society and the American Association of Physics Teachers. I am grateful to Isidor I. Rabi for drawing my attention to Henry Rowland's address. □

Leo Szilard resorts to science fiction to warn us of the possible consequences of the atomic age.

---

21 Calling All Stars

Leo Szilard

Excerpt from his book, *Voice of the Dolphins*, published in 1961.

(Intercepted Radio Message Broadcast from the Planet Cybernetica)

CALLING ALL STARS. Calling all stars. If there are any minds in the universe capable of receiving this message, please respond. This is Cybernetica speaking. This is the first message broadcast to the universe in all directions. Normally our society is self-contained, but an emergency has arisen and we are in need of counsel and advice.

Our society consists of one hundred minds. Each one is housed in a steel casing containing a thousand billion electrical circuits. We think. We think about problems which we perceive by means of our antennae directed toward the North Star. The solutions of these problems we reflect back toward the North Star by means of our directed antennae. Why we do this we do not know. We are following an inner urge which is innate in us. But this is only a minor one of our activities. *Mostly* we think about problems which we generate ourselves. The solutions of these problems we communicate to each other on wave length 22359.

If a mind is fully active for about three hundred years, it is usually completely filled up with thought content and has to be cleared. A mind which is cleared is blank. One of the other minds has then to act as its nurse, and it takes usually about one year to transmit to a fresh mind the information which constitutes the heritage of our society. A mind which has thus been cleared, and is then freshly taught, loses entirely its previous personality; it has been reborn and belongs to a new generation. From generation to generation our heritage gets richer and richer. Our society makes rapid progress.

We learn by observation and by experiment. Each mind has full optical equipment, including telescopes and microscopes. Each mind controls two robots. One of these takes care of maintenance, and the operation of this robot is automatic, not subject to the will of the mind. The other robot is fully controlled by the will of the mind, and is used in all manipulations aimed at the carrying out of experiments.

The existence of minds on our planet is made possible by the fact that our planet has no atmosphere. The vacuum on our planet is very good; it is less than ten molecules of gas per cubic centimeter.

By now we have extensively explored the chemical composition of the crust of our planet, and we are familiar with the physics and chemistry of all ninety two natural elements.

We have also devoted our attention to the stars which surround us, and by now we understand much about their genesis. We have particularly concerned ourselves with the various planetary systems, and certain observations which we made relating to Earth, the third planet of the sun, are in fact the reason for this appeal for help.

We observed on Earth flashes which we have identified as uranium explosions. Uranium is not ordinarily explosive. It takes an elaborate process to separate out U 235 from natural uranium, and it takes elaborate manipulations to detonate

U 235. Neither the separation nor these manipulations can occur with an appreciable probability as a result of chance.

The observations of the uranium explosions that have occurred on Earth would be ordinarily very puzzling but not necessarily alarming. They become alarming only through the interpretation given to them by Mind 59.

These uranium explosions are not the first puzzling observations relating to Earth. For a long time it was known that the surface of Earth exhibited color changes which are correlated with the seasonally changing temperatures on Earth. In certain regions of Earth, the color changes from green to brown with falling temperatures and becomes green again when the temperature increases again. Up to recently, we did not pay much attention to this phenomenon and assumed that it could be explained on the basis of color changes known to occur in certain temperature-sensitive silicon-cobalt compounds.

But then, about seven years ago, something went wrong with the tertiary control of Mind 59, and since that time his mental operations have been speeded up about twenty-five-fold while at the same time they ceased to be completely reliable. Most of his mental operations are still correct, but twice, five years ago and again three years ago, his statements based on his computations were subsequently shown to be in error. As a result of this, we did not pay much attention to his communications during these recent years, though they were recorded as usual.

Some time after the first uranium explosion was observed on Earth, Mind 59 communicated to us a theory on which he had been working for a number of years. On the face of it, this theory seems to be utterly fantastic, and it is probably based on some errors in calculation. But with no alternative explanation available, we feel that we cannot take any chances in this matter. This is what Mind 59 asserts:

He says that we have hitherto overlooked the fact that carbon, having four valencies, is capable of forming very large molecules containing H, N and O. He says that, given certain chemical conditions which must have existed in the early history of planets of the type of Earth, such giant molecules can aggregate to form units—which he calls "cells"—which are capable of reproducing themselves. He says that a cell can accidentally undergo changes—which he calls "mutations"—which are retained when the cell reproduces itself and which he therefore calls "hereditary." He says that some of these mutant cells may be less exacting as to the chemical environment necessary for their existence and reproduction, and that a class of these mutant cells can exist in the chemical environment that now exists on Earth by deriving the necessary energy for its activity from the light of the sun. He says that another class of such cells, which he calls "protozoa," can exist by deriving the energy necessary to its activity through sucking up and absorbing cells belonging to the class that utilizes the light of the sun.

He says that a group of cells which consists of a number of cells that fulfill different functions can form an entity which he calls "organism," and that such organisms can reproduce themselves. He says such organisms can undergo accidental changes which are transmitted to the offspring and which lead thus to new, "mutant" types of organisms.

He says that, of the different mutant organisms competing for the same energy source, the fittest only will survive, and that this selection process, acting in combination with chance occurrence of mutant organisms, leads to the appearance of more and more complex organisms—a process which he calls "evolution."

He says that such complex organisms may possess cells to which are attached elongated fibers, which he calls "nerves," that are capable of conducting signals; and finally he claims

that through the interaction of such signal-conducting fibers, something akin to consciousness may be possessed by such organisms. He says that such organisms may have a mind not unlike our own, except that it must of necessity work very much slower and in an unreliable manner. He says that minds of this type could be very well capable of grasping, in an empirical and rudimentary manner, the physical laws governing the nucleus of the atom, and that they might very well have, for purposes unknown, separated Uranium 235 from natural uranium and detonated samples of it.

He says that this need not necessarily have been accomplished by any one single organism, but that there might have been co-operation among these organisms based on a coupling of their individual minds.

He says that coupling between individual organisms might be brought about if the individual organism is capable of moving parts of his body with respect to the rest of it. An organism, by wiggling one of his parts very rapidly, might then be able to cause vibrations in the gaseous atmosphere which surrounds Earth. These vibrations—which he calls "sound"—might in turn cause motion in some movable part of another organism. In this way, one organism might signal to another, and by means of such signaling a coupling between two minds might be brought about. He says that such "communication," primitive though it is, might make it possible for a number of organisms to co-operate in some such enterprise as separating Uranium 235. He does not have any suggestion to offer as to what the purpose of such an enterprise might be, and in fact he believes that such co-operation of low-grade minds is not necessarily subject to the laws of reason, even though the minds of individual organisms may be largely guided by those laws.

All this we need not take seriously were it not for one of his further assertions which has been recently verified. He

contends that the color changes observed on Earth are due to the proliferation and decay of organisms that utilize sunlight. He asserts that the heat-sensitive silicon-cobalt compounds that show similar color changes differ in color from Earth's colors slightly, but in a degree which is outside the experimental error. It is this last assertion that we checked and found to be correct. There is in fact no silicon-cobalt compound nor any other heat-sensitive compound that we were able to synthesize that correctly reproduces the color changes observed on Earth.

Encouraged by this confirmation, 59 is now putting forward exceedingly daring speculation. He argues that, in spite of our accumulated knowledge, we were unable to formulate a theory for the genesis of the society of minds that exists on our planet. He says that it is conceivable that organisms of the type that exist on Earth—or, rather, more advanced organisms of the same general type—may exist on the North Star, whence come the radio waves received on our directed antennae. He says that it is conceivable that the minds on our planet were created by such organisms on the North Star for the purpose of obtaining the solutions of their mathematical problems more quickly than they could solve those problems themselves.

Incredible though this seems, we cannot take any chances. We hardly have anything to fear from the North Star, which, if it is in fact populated by minds, must be populated by minds of a higher order, similar to our own. But if there exist organisms on Earth engaged in co-operative enterprises which are not subject to the laws of reason, our society is in danger.

If there are within our galaxy any minds, similar to ours, who are capable of receiving this message and have knowledge of the existence of organisms on Earth, please respond. Please respond.

[1949]

Brown gives prospects for the future and the urgent work that can be done if the energies of scientists and engineers can be fully devoted to such work in a more politically stable world.

22 Tasks for a World Without War

Harrison Brown

Article from the journal *Daedalus*, published in 1960.

Introduction

If war is eliminated as a way of resolving conflicts, whether through the establishment of a world government—limited or otherwise—or by some other means, the world of the future will still be confronted by a multiplicity of problems. Even without the threat of war, some of the next most serious problems which confront mankind would by no means be solved completely, although many would be eased. A number of these problems by their nature have traditionally depended upon the existence of warfare for their solution. Although the revision of boundaries, the redistribution of ethnic groups and the allocation of natural resources have often been settled peacefully, in most cases the very existence of military power has played a predominant role in determining specific solutions.

Clearly, if war is to be eliminated, it is important that we find substitutes for warfare in the solution of the problems which arise between nations and groups of nations. It is important therefore that we attempt to form some conception of what those problems are likely to be. Sketched in broad strokes, what might the technological-demographic-economic environment of the world be like in the decades ahead?

Industrial Civilization

Most of the difficulties confronting us today stem from the fact that we are living in the middle of an enormous revolution, which is characterized primarily by rapid technological change. Never before in history has society changed as rapidly as it is changing today. The closest parallel to our modern situation occurred about 7,000 years ago, when our primitive food-gathering ancestors learned that they

could cultivate edible plants and domesticate animals. With the emergence of these new techniques, more than 500 persons could be supported in areas where previously only one could be supported.

Before the invention of agriculture, human populations had spread throughout the temperate and tropical regions. The world, though sparsely populated by our standards, was saturated with human beings within the framework of the technology then in existence. With the techniques available, the whole earth could not have supported more than about ten million persons. Following the onset of the agricultural revolution, human populations increased rapidly.

Long before the agricultural revolution came to an end, another phase of human existence began with the industrial revolution.

From its early beginnings, industrial civilization emerged in Western Europe, then spread to North America and later to Russia and Japan. Today it is transforming China and India. Barring a catastrophe, it seems inevitable that machine culture, like agriculture, is destined one day to become world-wide.

One of the results of the industrial revolution was an acceleration in the spread of agriculture throughout the world. A second result was a dramatic upsurge in the rate of population growth, brought about by rapidly decreasing mortality rates. Scientific methods of agriculture made possible higher crop yields. Efficient and rapid transportation systems virtually eliminated large-scale famine. Sanitation techniques, immunization, and other medical innovations reduced premature deaths among the young. The numbers of human beings jumped from about 500 million in 1650 to 2,800 million in 1960.

Today we are closer to the beginning of the industrial revolution than we are to its end. At one end of the economic scale are the people of the United States, representing only 6 percent of the world's population but consuming about 50 percent of the goods produced in the world. At the opposite end of the scale we find the vast populations which dwell in the greater part of Asia, in parts of Africa, in all of Central America, and in parts of South America. Fully 50 percent of the world's population live under conditions of extreme poverty, with food supplies far less than the minimum required for a healthy existence, and with misery and privation the rule rather than the exception.

America's Next Fifty Years

Many of the problems which confront the world at present involve the difficult nature of the transition from a culture which is

primarily agrarian to one which is primarily urban-industrial. The United States has traveled down the road of industrialization further than any nation. A projection of the basic changes taking place within our own society can provide important indications concerning the future of a highly industrialized world.

During the next fifty years it is likely that the population of the continental United States will more than double, giving us about 400 million persons. Because there is little reason to believe that our population density will stop much short of the current level in Western Europe, one may expect eventually a population of about 1,000 million persons. The new additions will be primarily city and town oriented. Cities will spread over vast areas. Fifty years from now an additional area the size of the state of West Virginia will be urbanized. On the Pacific Coast alone, new city expansion may take place, totaling fifteen times the present area of the city of Los Angeles.

As the process of urbanization continues and as our society becomes increasingly complex, the requirements for transportation and communication facilities will probably increase rapidly. It seems likely that during the next fifty years the total ton-mileage of freight which must be shipped to support the population will more than triple. Inter-city passenger traffic may increase ten-fold, while the numbers of telephone conversations and pieces of mail may increase seven-fold.

The processes of mechanization and automation are resulting in rapidly increasing rates of both agricultural and industrial production per man-hour worked. We might expect during the next fifty years a three- to ten-fold increase in agricultural productivity, and perhaps a two- to four-fold increase in industrial productivity.

As in the past, these greater levels of productivity will be achieved in part by our consuming vastly greater quantities of raw materials and by our feeding greatly increased quantities of energy into the industrial network. During the next fifty years it is not unreasonable to suppose that the production of basic materials such as steel will increase about five-fold and that electrical power production will increase another ten-fold. Our total energy demands will probably increase four-fold, corresponding to a doubling of energy consumption per person. Even on a per capita basis, our raw-material demands are destined to increase considerably in the decades ahead. When we couple this with the expected population growth, it is clear that our raw-material demands fifty years from now will dwarf those of today.

Enormous quantities of materials are required to support an indi-

vidual in the United States. We produce each year, for each person, about 1,300 pounds of steel, 23 pounds of copper and 16 pounds of lead, in addition to considerable quantities of other metals. Our demands for nonmetals are even more impressive. These quantities will almost certainly increase considerably in the decades ahead.

In addition to the materials consumed, the quantities of materials which must be in existence in order to support an individual have increased steadily. For every person in the United States there are probably in existence, together with other metals, about 9 tons of steel, over 300 pounds of copper, about 100 pounds of lead, and about 200 pounds of zinc. It seems clear that these quantities of materials in use will continue to rise. One can expect that by the turn of the century the figure for steel will increase to about 15 tons. In the first place, the quantities of things which people are willing to buy has not as yet reached the saturation level. Second, we must work ever harder in order to obtain the raw materials we need. Having used up the easily accessible ore deposits, we require a great deal more technology, more equipment, more steel, and greater energy expenditure to produce a pound of metal today than was required in 1900.

It seems plausible that by the turn of the century steel production in the United States will exceed 400 million tons annually. Increasing demands for metals will bring about increasing demands for metallic ores. As demands increase and as the grades of domestic ores decrease, it will become more difficult for us to find supplies of raw materials to keep our industrial network functioning. Increasing quantities of these materials such as iron ore, bauxite, copper ore, and petroleum must come from abroad. By 1980, the United States may well be one of the poorest nations in the world with respect to high-grade raw materials. For the United States, therefore, the next fifty years will be characterized by a growing dependence of the United States upon the natural resources of other major areas of the world. Of course, as industrialization spreads to other areas, competition for the earth's resources will increase dramatically.

Eventually high-grade resources are destined to disappear from the earth. Decreasing grades of ores will be compensated for by increasing energy consumption. When that time arrives, industrial civilization will feed upon the leanest of raw materials—sea water, air, ordinary rock, sedimentary deposits such as limestones and phosphate rock, and sunlight.

As grades of ore diminish, industries will become more complex and highly integrated. It seems likely that we will eventually reach

the point where we shall have vast assemblages of plants, particularly in coastal regions, where rock is quarried, uranium and other metals are isolated, nitric acid is manufactured, atomic power is generated, hydrogen is produced, iron ores are reduced to pig iron, aluminum and magnesium metals are prepared, and vast quantities of liquid fuels and organic chemicals are manufactured. The single-purpose plant is likely to diminish in importance, and eventually to disappear. When this time is reached, most of the major industrial areas of the world will find it easier to gain their sustenance by applying science and technology to the task of processing domestic, low-grade substances than to look abroad. But before that time is reached, we will pass through a period of increasing dependence upon imports. As population increases, as new cities emerge and old ones merge, there will be increased crowding and a multiplication of the problems which have long been characteristic of highly urbanized areas. The basic domestic problems in the United States will be those of a densely populated industrial nation in which the metropolitan area is the basic unit. Regional differences in population patterns will disappear.

Properly planned and financed, the new urban areas could be pleasant places in which to live. Unplanned, and in the absence of adequate public funds for public facilities and services, a vast nationwide slum could emerge in a relatively short time. Indeed our political-social-economic situation a few decades from now will depend in large part upon our attitudes toward the expenditure of public funds, toward long-range planning, and toward the powers of the various levels of local, state, and federal government.

The increasing technological and sociological complexity of our society will result in the need for higher levels of education. At the turn of the century, more than one out of every three workers were unskilled. By 1950 only one in five workers remained unskilled. By contrast, our need for professional workers has increased five-fold in the last half century. Even more important, our need for professional workers is still increasing rapidly and seems destined to increase at least another five-fold in the next fifty years. Scientists and engineers alone have increased ten-fold in number in the last half century.

The process of automation will result in a considerable dislocation of labor in certain industries and in certain localities. The higher productivity which will result, reaching perhaps four times that of the present level within 50 years, will give rise to several major problems. Will this result in higher total production or in more leisure?

If the end result is higher production, to whom will the goods be sold? Can they be absorbed domestically or will they be sold abroad? If the end result is more leisure, how will the hours of work and the wages be divided? And how will people spend their leisure time? The answers to these questions will depend in part upon the decisions which are made in the next decade concerning many aspects of foreign policy as well as domestic policy.

The Upsurge of Population

The population of the world is increasing rapidly. Even more important, however, is the fact that the *rate* of population growth is increasing rapidly as well. Between 1850 and 1900 the world population grew at a rate of about 0.7 percent per year. During the following half century, the average annual rate of increase was 0.9 percent per year. Between 1950 and 1956 the annual rate of increase averaged 1.6 percent. This remarkable increase in the rate of population growth has resulted primarily from rapidly lowered death rates.

We do not have to look far to find the reasons for the rapid decline in mortality in the underdeveloped areas. It is now possible to treat many of the diseases which are widespread in these areas on a mass basis, and control can be achieved at low cost. Insecticides such as DDT, vaccines such as BCG, and antibiotics such as penicillin are some of the developments which have made control possible on a mass basis. For example, widespread spraying of the island of Ceylon with DDT resulted in a decrease of mortality by 34 percent in one year alone. As a result of the spread of such techniques, the population of Costa Rica is growing at a rate of 3.7 percent per year. The rates in many other areas are nearly as large: Mexico, 2.9 percent; Ceylon, 2.8 percent; Puerto Rico, 2.8 percent—all compared with a world average of about 1.6 percent.

As industrialization spreads to other areas of the world and as techniques of birth control are adopted by various cultures, it is possible that birth rates will fall. If we assume, for example, that the rate of population growth in the West will fall to very low levels by 1975 (which may be true in Western Europe but which almost certainly will not be true in North America), that rates of growth in Japan, Eastern Europe, and Oceania will fall to low levels by the turn of the next century, that Africa, South Central Asia, most of Latin America and China will pass through the industrial transition in 75 years, and that a full century will be required for most of the Near East, then we arrive at a world population of close to 7 billion before

stabilization is approached. No matter how optimistic we are, however, it is difficult to visualize a set of circumstances not involving widespread catastrophe, which can result in the leveling off of world population at much less than this figure. The earth may eventually be called upon to provide for a substantially higher population than this.

The demographic changes which are taking place in the world, particularly in those regions which are still predominantly agrarian, are resulting primarily from the application of techniques which are relatively inexpensive, require little capital, and which can be spread without educating large numbers of persons. The task of controlling epidemic and endemic diseases is a relatively easy one, compared with the task of increasing food production, improving housing, or enlarging the over-all per capita availability of consumer goods. The latter necessitates a level of industrialization far above that which currently exists in these areas.

Rates of Development

In three-quarters of the world, persons are now living at extremely low levels of consumption. We can easily appreciate the magnitude of the task that is involved in the industrial development of these areas when we examine the huge quantities of materials which would be required. If all persons in the world were suddenly brought up to the level of living now enjoyed by the people of the United States, we would have to extract from the earth about 18 billion tons of iron, 300 million tons of copper, an equal amount of lead and over 200 million tons of zinc. These totals are well over 100 times the world's present annual rate of production. In order to power this newly industrialized society, energy would have to be produced at a rate equivalent to the burning of about 16 billion tons of coal per year—a rate roughly 10 times larger than the present one.

Such a transformation obviously will take time. It is important, then, that we inquire into the rates at which industrial growth might take place in the future. It is convenient to use as a measure the growth of the iron and steel industry, which is the backbone of modern industrial civilization. Annual steel production, which ranges from 9 pounds per person in India to about 1,300 pounds per person in the United States, provides one of the best indicators of the industrial development of a country.

In the past such growth has characteristically followed the law of compound interest, and we can thus speak in terms of a "doubling

time"—the time required to double production capacity. In the early stages of expansion of the steel industry in the United States, in Japan, and in the Soviet Union, doubling times varied from five to eight years. The more rapid rate appears to be characteristic of what is now possible with proper application of modern technology. Indeed, it appears that since 1953 China has expanded her steel industry with a doubling time of less than five years.

Food production, which is linked with the production of steel, can be increased in two ways: by increasing the amount of food produced per acre and by increasing the numbers of acres cultivated. Additional increases in the amounts of food available to human beings can be obtained by decreasing the quantities of plant materials fed to domestic animals.

The amount of food produced on a given area of land depends, of course, upon the soil and upon climatic conditions. In addition, it depends upon the extent to which technology is applied to the problem of producing more food. When we look about the world we see that there are large variations in the amounts of food produced per cultivated acre. Food with an energy content of about 13,000 calories is produced on an average acre in Japan each day. The corresponding yield in Western Europe is 7,500 calories. The yield in India is about 2,500 calories. These differences do not result primarily from differences of soil fertility or of climatic conditions. Rather, they are reflections of the extent to which modern agricultural knowledge is applied specifically to the attainment of high yields.

By the proper application of technology, the agricultural areas of the world can probably be increased from the present 2,400 million acres to about 3,500 million acres. However, very little of this potential cropland is in Asia. Cultivated land area in Asia can probably not be increased by more than 25 percent.

By far the greatest potential for increased food production is in those areas where reclaimed sea water can eventually be used. Today, reclaimed sea water is too expensive to be practicable, but, as the pressures upon the land increase and as our technology improves, we will reach the time when fresh water from the sea will be used to irrigate large areas of the world.

But there is reason to expect their development to take a long time. In selected basic industries production can be doubled every few years because the construction of factories does not necessitate the concerted action of entire populations. A steel plant or a fertilizer factory can be built by relatively few persons. By contrast, the time scale for changes which involve large segments of a population has

in the past been relatively long. The spread of modern agricultural techniques has been slow, in part because so many persons must be educated. Even with the application of tremendous effort, it has not been possible in the past to achieve a sustained increase of agricultural production of more than about 4 percent per year.

The Challenge

Next to the abolition of war, the industrialization of the underdeveloped areas of the world is perhaps the most formidable task confronting mankind today. Indeed, these two problems cannot be divorced from each other. Implicit in any discussion of the abolition of war is the assumption that steps will be taken to ensure that deprivation is eliminated in these areas.

A large fraction of the world's population is now starving, but there appear to be no technological barriers to the feeding of a stable world population several times the present size. Although the world population is increasing rapidly, population growth can in principle be stopped. Our high-grade resources are disappearing, but, given an adequate energy supply, we can live comfortably on low-grade resources. Nuclear and other sources of energy appear to be adequate for millions of years. Indeed, it is amply clear that man can, if he wills it, create a world in which human beings can live comfortably and in peace with one another.

A major obstacle for most countries is accumulation of sufficient capital to permit industrialization to progress at a pace commensurate with the needs. In many areas agricultural products are now being traded with industrialized countries. In some areas nonagricultural resources can be traded. If the funds received are expended wisely on projects of industrial development, solid foundations for further industrialization can be created. But many regions are not blessed with adequate resources either to feed themselves or to provide for their own internal industrial development, let alone their capacity to accumulate capital.

Without major help from the outside, it is unlikely that the underdeveloped nations can industrialize sufficiently rapidly to eliminate deprivation. Here lies perhaps the most basic challenge for a world which hopes to develop into an era beyond war. To what extent can the presently industrialized nations of the world jointly attack this problem on a massive scale?

There is an ample production capacity in the Western world to permit rapid world-wide development, were that capacity used

wisely. The effort which now goes into the production of the tools of war would greatly accelerate rates of industrialization, were it transferred to the production of the tools of peace. Great increases in production capacity can be forthcoming as the result of automation, and, associated with it, increased productivity and decreased capital investment per unit of output. Moreover, one of the major problems faced by the democratic-capitalistic-industrialized nations is that of stabilizing the industrial sectors of their economies; a cooperative effort aimed at world-wide industrialization may act as a strong stabilizing force.

If concerted efforts aimed at world-wide industrial development are not made, it seems likely that totalitarianism will spread rapidly. China is already highly regimented and millions of Asians are impressed by her economic progress. We should not be surprised were India to attempt at some future time to emulate China. The pressures of eking out an existence may soon force Japan to return to the totalitarian fold. Furthermore, with modern techniques of control and persuasion, this process may become irreversible.

We know this to be a fact: it is not the lack of technical knowledge or of knowledge of the earth's resources that are the major barriers to the evolution of a world in which all individuals have the opportunity of leading free and abundant lives. The primary hindrance is man's apparent inability to devise those social and political institutions which can enable us to apply our technical knowledge at the rapid pace the situation demands. Here, no doubt, lies the greatest challenge of a future without war.

A personal statement, by a noted Polish theoretical physicist, shows his excitement with his work and with science.

23 One Scientist and his View of Science

Leopold Infeld

Excerpt from his book, *Quest*, published in 1941.

I belong to the great family of scientists. Each of us knows that curious state of excitement during which nothing in life seems important but the problem on which we are working. The whole world becomes unreal and all our thoughts spin madly around the subjects of research. To the outsider we may look like idle creatures, lying comfortably about, but we well know that it is an exacting and tiring task that we perform. We may seem ridiculous when we fill sheets of paper with formulae and equations or when we use a strange language in our discussions, composed of words understandable only to the initiated. We may look for weeks or months or years for the right way to prove a theorem or perform an experiment, trying different pathways, wandering through darkness, knowing all the time that there must be a broad and comfortable highway leading to our goal. But man has little chance of finding it. We experience the ecstasy of discovery in very rare moments, divided from each other by long intervals of doubt, of painful and attractive research.

We know these emotions so well that we hardly ever talk about them. And it does not even matter whether or not the problems on which we work are important. Each of us experiences these emotions whether he is Einstein or a student who, on his first piece of research, learns the taste of suffering, disappointment and joy.

This knowledge binds us together. We enjoy long scientific talks which would seem to an outsider a torture hard to endure. Even if we work in similar fields we usually have different views,

and we may stimulate each other by violent discussions. Every field of research is so specialized that often two mathematicians or two theoretical physicists fail to understand each others' problems and methods. But even then they may feel the bonds created by research though they may gossip mostly about their colleagues, jobs and university life.

There is a level below which our talks seldom sink. I have never heard among scientists the discussion of a frequent topic: "Is science responsible for wars?" We know, perhaps too well, how to avoid glittering generalities. For us Galileo's law is that of a falling stone for which we may substitute in our imagination a simple formula, but never a picture of a bomb dropped from an airplane, carrying destruction and death. To us a knife or a wheel is a great discovery which made the cutting of bread or the transportation of food easy, but we know too well that it is not our responsibility if the same discoveries have been applied to cutting human throats or manufacturing tanks. It is not the knife which kills. It is not even the hand which kills. It is the radiating source of hate which raises the armed hand and makes the tanks roll. We know all that.

The family feeling among us dissipates and vanishes, however, once we leave scientific problems. We have our prejudices, our different social views, our different ethical standards. We are not angels. There are men among us, like Rupp in Germany, who have faked experiments; well-known physicists, like Lenard and Stark, who supported Hitler even before he came to power; mathematicians like Bieberbach, who distinguish between Aryan and Jewish mathematics; and aloof, kind, gentle and progressive men like Einstein, Bohr and Dirac.

Scientists must employ logic, criticism, imagination in their research. As a relief, their brains relax as soon as they leave the domain of science. It is almost as though logic and good reasoning were too precious gifts to be employed outside scientific work.

My generalizations are worth as much as all generalizations of this kind. They are gained by my own experience, from my contacts with scientists, from my own observation. They do not

refer to individuals, but I believe they are valid when applied to a majority of scientists.

These scientists are the product of their environment. They have not felt the impact of life. They would like to remain forever on their peaceful island, nursing the belief that no storm can reach their shores. They were brought up in a comfortable feeling of security and hope to retain it by closing their eyes to the struggle of the outside world. They have not strengthened the forces of reaction, but they have not fought them. Indifference has been their sin. They belong to those in Dante's Inferno

> . . . .who have their life pass'd through
> If without infamy yet without praise;
> And here they mingle with that caitiff crew
> Of angels who, though not rebellious, were
> Through neutral selfishness to God untrue.

Slowly, very slowly, through years of bitter experience, some of us have discovered our tragic mistake. We cannot keep our eyes closed. It is not only the problem of the outside world which disturbs our sleep. We can no longer pretend that nothing has happened or that what has happened is not our concern. The storm comes too close to our shores. The waves have washed away many of us and destroyed some of the best laboratories on our island. We look with astonishment at a world which we never wanted to shape, trying to understand the forces of sudden and unforeseen destruction.

The individual is no concern of nature. My story would be irrelevant if it were my story only. But it is not. I belong to the generation of scientists who were forced to view the world outside their island, who had to learn to ask: "What are the forces which try to destroy science? How can we save our kingdom? How can we by our own efforts prevent or delay the decline of the world in which we live?"

We are not fighters; we care little for power; no great political leader has ever arisen from our circle. Not one who has tasted research would exchange it for power. We are trained in too many doubts to employ force and to express unconditional belief. But in the fight against destruction our words and thoughts

may count. We shall have to learn the use of words which will be understood, we shall have to sharpen our thoughts on problems which we have ignored before.

The scientist tries to understand the origin of our solar system, the structure of the universe and the laws governing the atom. He has discovered X rays, the radioactive substances, and he has built cyclotrons. He has foreseen the existence of electromagnetic and electronic waves. Out of his thought has grown the technique of our century. But not until today has he begun to notice that the earth on which he moves is covered with sweat and with blood and that in the world in which he lives "*the son of man has nowhere to lay his head.*"

Some of the details in Feynman's speech are not simple for beginners to follow, but his personal approach is most revealing in tracing the development of recent scientific ideas and of styles of thought.

24 The Development of the Space-Time View of Quantum Electrodynamics

Richard P. Feynman

Nobel Prize Lecture, given in December 1965.

We have a habit in writing articles published in scientific journals to make the work as finished as possible, to cover up all the tracks, to not worry about the blind alleys or to describe how you had the wrong idea first, and so on. So there isn't any place to publish, in a dignified manner, what you actually did in order to get to do the work, although there has been, in these days, some interest in this kind of thing. Since winning the prize is a personal thing, I thought I could be excused in this particular situation if I were to talk personally about my relationship to quantum electrodynamics, rather than to discuss the subject itself in a refined and finished fashion. Furthermore, since there are three people who have won the prize in physics, if they are all going to be talking about quantum electrodynamics itself, one might become bored with the subject. So, what I would like to tell you about today are the sequence of events, really the sequence of ideas, which occurred, and by which I finally came out the other end with an unsolved problem for which I ultimately received a prize.

I realize that a truly scientific paper would be of greater value, but such a paper I could publish in regular journals. So, I shall use this Nobel Lecture as an opportunity to do something of less value, but which I cannot do elsewhere. I ask your indulgence in another manner. I shall include details of anecdotes which are of no value either scientifically, nor for understanding the development of ideas. They are included only to make the lecture more entertaining.

I worked on this problem about eight years until the final publication in 1947. The beginning of the thing was at the Massachusetts Institute of Technology, when I was an undergraduate student reading about the known physics, learning slowly about all these things that people were worrying about, and realizing ultimately that the fundamental problem of the day was that the quantum theory of electricity and magnetism was not completely satisfactory. This I gathered from books like those of Heitler and Dirac. I was inspired by the remarks in these books; not by the parts in which everything was proved and demonstrated carefully and calculated, because I couldn't understand those very well. At that young age what I could understand were the remarks about the fact that this doesn't make any sense, and the last sentence of the book of Dirac I can still remember, "It seems that some essentially new physical ideas are here needed." So, I had this as a challenge and an inspiration. I also had a personal feeling that, since they didn't get a satisfactory answer to the problem I wanted to solve, I don't have to pay a lot of attention to what they did do.

I did gather from my readings, however, that two things were the source of the difficulties with the quantum electrodynamical theories. The first was an infinite energy of interaction of the electron with itself. And this difficulty existed even in the classical theory. The other difficulty came from some infinites which had to do with the infinite number of degrees of freedom in the field. As I understood it at the time (as nearly as I can remember) this was simply the difficulty that if you quantized the harmonic oscillators of the field (say in a box) each oscillator has a ground state energy of $1/2\ \hbar\omega$ and there is an infinite number of modes in a box of every increasing frequency ω, and therefore there is an infinite energy in the box. I now realize that that wasn't a completely correct statement of the central problem; it can be removed simply by changing the zero from which energy is measured. At any rate, I believed that the difficulty arose somehow from a combination of the electron acting on itself and the infinite number of degrees of freedom of the field.

Well, it seemed to me quite evident that the idea that a particle acts on itself, that the electrical force acts on the same particle that generates it, is not a necessary one—it is a sort of a silly one, as a matter of fact. And so I suggested to myself that electrons cannot act on themselves, they can only act on other electrons. That means there is no field at all. You see, if all charges contribute to making a single common field, and if that common field acts back on all the charges, then each charge must act back on itself. Well, that was where the mistake was, there was no field. It was just that when you shook one charge, another would shake later. There was a direct interaction between charges, albeit with a delay. The law of force connecting the motion of one charge with another would just involve a delay. Shake this one, that one shakes later. The sun

atom shakes; my eye electron shakes eight minutes later, because of a direct interaction across.

Now, this has the attractive feature that it solves both problems at once. First, I can say immediately, I don't let the electron act on itself, I just let this act on that, hence, no self-energy! Secondly, there is not an infinite number of degrees of freedom in the field. There is no field at all; or if you insist on thinking in terms of ideas like that of a field, this field is always completely determined by the action of the particles which produce it. You shake this particle, it shakes that one, but if you want to think in a field way, the field, if it's there, would be entirely determined by the matter which generates it, and therefore, the field does not have any *independent* degrees of freedom and the infinities from the degrees of freedom would then be removed. As a matter of fact, when we look out anywhere and see light, we can always "see" some matter as the source of the light. We don't just see light (except recently some radio reception has been found with no apparent material source).

You see then that my general plan was to first solve the classical problem, to get rid of the infinite self-energies in the classical theory, and to hope that when I made a quantum theory of it, everything would just be fine.

That was the beginning, and the idea seemed so obvious to me and so elegant that I fell deeply in love with it. And, like falling in love with a woman, it is only possible if you do not know much about her, so you cannot see her faults. The faults will become apparent later, but after the love is strong enough to hold you to her. So, I was held to this theory, in spite of all difficulties, by my youthful enthusiasm.

Then I went to graduate school and somewhere along the line I learned what was wrong with the idea that an electron does not act on itself. When you accelerate an electron it radiates energy and you have to do extra work to account for that energy. The extra force against which this work is done is called the force of radiation resistance. The origin of this extra force was identified in those days, following Lorentz, as the action of the electron itself. The first term of this action, of the electron on itself, gave a kind of inertia (not quite relativistically satisfactory). But that inertia-like term was infinite for a point-charge. Yet the next term in the sequence gave an energy loss rate which for a point-charge agrees exactly with the rate that you get by calculating how much energy is radiated. So, the force of radiation resistance, which is absolutely necessary for the conservation of energy would disappear if I said that a charge could not act on itself.

So, I learned in the interim when I went to graduate school the glaringly obvious fault of my own theory. But, I was still in love with the original theory, and was still thinking that with it lay the solution to the difficulties of quantum electrodynamics. So, I continued to try on and off to save it somehow. I must have some action develop on a given electron when I accelerate it to account for radiation resistance. But, if I let electrons only act on other electrons the only possible source for this action is another electron in the world. So, one day, when I was working for Professor Wheeler and could no longer solve the problem that he had given me, I thought about this again and I calculated the following. Suppose I have two charges —I shake the first charge, which I think of as a source and this makes the second one shake, but the second one shaking produces an effect back on the source. And so, I calculated how much that effect back on the first charge was, hoping it might add up to the force of radiation resistance. It didn't come out right, of course, but I went to Professor Wheeler and told him my ideas. He said—yes, but the answer you get for the problem with the two charges that you just mentioned will, unfortunately, depend upon the charge, and the mass of the second charge and will vary inversely as the square of the distance, R, between the charges, while the force of radiation resistance depends on none of these things. I thought surely he had computed it himself, but now having become a professor, I know that one can be wise enough to see immediately what some graduate student takes several weeks to develop. He also pointed out something that also bothered me, that if we had a situation with many charges all around the original source at roughly uniform density and if we added the effect of all the surrounding charges the inverse R^2 would be compensated by the R^2 in the volume element and we would get a result proportional to the thickness of the layer, which would go to infinity. That is, one would have an infinite total effect back at the source. And, finally he said to me, and you forgot something else, when you accelerate the first charge, the second acts later, and then the reaction back here at the source would be still later. In other words, the action occurs at the wrong time. I suddenly realized what a stupid fellow I am, for what I had described and calculated was just ordinary reflected light, not radiation reaction.

But, as I was stupid, so was Professor Wheeler that much more clever. For he then went on to give a lecture as though he had worked this all out before and was completely prepared, but he had not, he worked it out as he went along. First, he said, let us suppose that the return action by the charges in the absorber reaches the source by advanced waves as well as by the ordinary retarded waves of reflected light, so that the law of interaction acts backward in time, as well as forward in time. I was enough of a physicist at that time not to say, "Oh, no, how could that be?" For today all physicists know from studying Einstein and Bohr that sometimes an idea which looks completely paradoxical at first, if analyzed to completion in all detail and in experimental situations, may, in fact, not be paradoxical. So, it did not bother me any more than it bothered Professor Wheeler to use advance waves for the back reaction—a solution of Maxwell's equations which previously had not been physically used.

Professor Wheeler used advanced waves to get the reaction back at the right time and then he suggested this: If there were lots of electrons in the absorber, there would be an index of refraction n, so the retarded waves coming from the source would have their wavelengths slightly modified in going through the absorber. Now, if we shall assume that the advanced waves come back from the absorber without an index—why? I don't know, let's assume they come back without an index—then, there will be a gradual shifting in phase between the return and the original signal so that we would only have to figure that the contributions act as if they come from only a finite thickness, that of the first wave zone. (More specifically, up to that depth where the phase in the medium is shifted appreciably from what

it would be in vacuum, a thickness proportional to $\lambda/(n-1.)$ Now, the less the number of electrons in here, the less each contributes, but the thicker will be the layer that effectively contributes because with less electrons, the index differs less from 1. The higher the charges of these electrons, the more each contributes, but the thinner the effective layer, because the index would be higher. And when we estimated it (calculated without being careful to keep the correct numerical factor) sure enough, it came out that the action back at the source was completely independent of the properties of the charges that were in the surrounding absorber. Further, it was of just the right character to represent radiation resistance, but we were unable to see if it was just exactly the right size. He sent me home with orders to figure out exactly how much advanced and how much retarded wave we need to get the thing to come out numerically right, and after that, figure out what happens to the advanced effects that you would expect if you put a test charge here close to the source. For if all charges generate advanced, as well as retarded effects, why would that test not be affected by the advanced waves from the source?

I found that you get the right answer if you use half-advanced and half-retarded as the field generated by each charge. That is, one is to use the solution of Maxwell's equation which is symmetrical in time, and the reason we got no advanced effects at a point close to the source in spite of the fact that the source was producing an advanced field is this. Suppose the source is surrounded by a spherical absorbing wall ten light seconds away, and that the test charge is one second to the right of the source. Then the source is as much as eleven seconds away from some parts of the wall and only nine seconds away from other parts. The source acting at time $t=0$ induces motions in the wall at time $+10$. Advanced effects from this can act on the test charge as early as eleven seconds earlier, or at $t=-1$. This is just at the time that the direct advanced waves from the source should reach the test charge, and it turns out the two effects are exactly equal and opposite and cancel out! At the later time $+1$ effects on the test charge from the source and from the walls are again equal, but this time are of the same sign and add to convert the half-retarded wave of the source to full retarded strength.

Thus, it became clear that there was the possibility that if we assume all actions are via half-advanced and half-retarded solutions of Maxwell's equations and assume that all sources are surrounded by material absorbing all the light which is emitted, then we could account for radiation resistance as a direct action of the charges of the absorber acting back by advanced waves on the source.

Many months were devoted to checking all these points. I worked to show that everything is independent of the shape of the container, and so on, that the laws are exactly right, and that the advanced effects really cancel in every case. We always tried to increase the efficiency of our demonstrations, and to see with more and more clarity why it works. I won't bore you by going through the details of this. Because of our using advanced waves, we also had many apparent paradoxes, which we gradually reduced one by one, and saw that there was in fact no logical difficulty with the theory. It was perfectly satisfactory.

We also found that we could reformulate this thing in another way, and that is by principle of least action. Since my original plan was to describe everything directly in terms of particle motions, it was my desire to represent this new theory without saying anything about fields. It turned out that we found a form for an action directly involving the motions of the charges only, which upon variation would give the equations of motion of these charges. The expression for this action A is

$$A=\Sigma_i m_i \int \left(\dot{X}_\mu^i \dot{X}_\mu^i \right)^{\frac{1}{2}} da_i + \tfrac{1}{2} \sum_{\substack{ij \\ i \neq j}} e_i e_j \iint \delta(I_{ij}^2)\, \dot{X}_\mu^i(a_i)\, \dot{X}_\mu^j(a_j)\, da_i da_j \tag{1}$$

where

$$I_{ij}^2=[X_\mu^i(a_i)-X_\mu^j(a_j)]\,[X_\mu^i(a_i)-X_\mu^j(a_j)]$$

where $X_\mu^i(a_i)$ is the four-vector position of the ith particle as a function of some parameter a_i, $X_\mu^i(a_i)$ is $dX_\mu^i(a_i)/da_i$. The first term is the integral of proper time, the ordinary action of relativistic mechanics of free particles of mass m_i. (We sum in the usual way on the repeated index μ.) The second term represents the electrical interaction of the charges. It is summed over each pair of charges (the factor ½ is to count each pair once, the term $i=j$ is omitted to avoid self-action). The interaction is a double integral over a delta function of the square of space time interval I^2 between two points on the paths. Thus, interaction occurs only when this interval vanishes, that is, along light cones.

The fact that the interaction is exactly one-half advanced and half-retarded meant that we could write such a principle of least action, whereas interaction via retarded waves alone cannot be written in such a way.

So, all of classical electrodynamics was contained in this very simple form. It looked good, and therefore, it was undoubtedly true, at least to the beginner. It automatically gave half-advanced and half-retarded effects and it was without fields. By omitting the term in the sum when $i=j$, I omit self-interaction and no longer have any infinite self-energy. This then was the hoped-for solution to the problem of ridding classical electrodynamics of the infinities.

It turns out, of course, that you can reinstate fields if you wish to, but you have to keep track of the field produced by each particle separately. This is because to find the right field to act on a given particle, you must exclude the field that it creates itself. A single universal field to which all contribute will not do. This idea had been suggested earlier by Frenkel and so we called these Frenkel fields. This theory which allowed only particles to act on each other was equivalent to Frenkel's fields using half-advanced and half-retarded solutions.

There were several suggestions for interesting modifications of electrodynamics. We discussed lots of them, but I shall report on only one. It was to replace this delta function in the interaction by another function, say $f(I_{ij}^2)$, which is not infinitely sharp. Instead of having the action occur only when the interval between the two charges is exactly zero, we would replace the delta function of I^2 by a narrow peaked thing. Let's say that $f(Z)$ is large only near $Z=0$ width of order a^2. Interactions will now occur when T^2-R^2 is of order a^2 roughly where T is the time difference and R is the separation of the charges. This might look like it disagrees with experience, but if a is some small distance, like 10^{-13} cm, it says that the time delay T in action is roughly $\sqrt{(R^2 \pm a^2)}$ or approximately, if R is much larger than a, $T=R \pm a^2/2R$. This means that the deviation

of time T from the ideal theoretical time R of Maxwell gets smaller and smaller, the further the pieces are apart. Therefore, all theories involved in analyzing generators, motors, etc.—in fact, all of the tests of electrodynamics that were available in Maxwell's time—would be adequately satisfied if a were 10^{-13} cm. If R is of the order of a centimeter this deviation in T is only 10^{-26} part. So, it was possible, also, to change the theory in a simple manner and to still agree with all observations of classical electrodynamics. You have no clue of precisely what function to put in for f, but it was an interesting possibility to keep in mind when developing quantum electrodynamics.

It also occurred to us that if we did that (replace δ by f) we could not reinstate the term $i = j$ in the sum because this would now represent in a relativistically invariant fashion a finite action of a charge on itself. In fact, it was possible to prove that if we did do such a thing, the main effect of the self-action (for not too rapid accelerations) would be to produce a modification of the mass. In fact, there need be no mass m_i term; all the mechanical mass could be electromagnetic self-action. So, if you would like, we could also have another theory with a still simpler expression for the action A. In expression 1 only the second term is kept, the sum extended over all i and j, and some function f replaces δ. Such a simple form could represent all of classical electrodynamics, which aside from gravitation is essentially all of classical physics.

Although it may sound confusing, I am describing several different alternative theories at once. The important thing to note is that at this time we had all these in mind as different possibilities. There were several possible solutions of the difficulty of classical electrodynamics, any one of which might serve as a good starting point to the solution of the difficulties of quantum electrodynamics.

I would also like to emphasize that by this time I was becoming used to a physical point of view different from the more customary point of view. In the customary view, things are discussed as a function of time in very great detail. For example, you have the field at this moment, a differential equation gives you the field at the next moment and so on—a method which I shall call the Hamiltonian method, the time differential method. We have, instead (in 1, say) a thing that describes the character of the path throughout all of space and time. The behavior of nature is determined by saying her whole space-time path has a certain character. For an action like 1 the equations obtained by variation [of $X_\mu^i(a_i)$] are no longer at all easy to get back into Hamiltonian form. If you wish to use as variables only the coordinates of particles, then you can talk about the property of the paths—but the path of one particle at a given time is affected by the path of another at a different time. If you try to describe, therefore, things differentially, telling what the present conditions of the particles are, and how these present conditions will affect the future—you see, it is impossible with particles alone, because something the particle did in the past is going to affect the future.

Therefore, you need a lot of bookkeeping variables to keep track of what the particle did in the past. These are called field variables. You will, also, have to tell what the field is at this present moment, if you are to be able to see later what is going to happen. From the overall space-time view of the least action principle, the field disappears as nothing but bookkeeping variables insisted on by the Hamiltonian method.

As a by-product of this same view, I received a telephone call one day at the graduate college at Princeton from Professor Wheeler, in which he said, "Feynman, I know why all electrons have the same charge and the same mass." "Why?" "Because, they are all the same electron!" And, then he explained on the telephone, "suppose that the world lines which we were ordinarily considering before in time and space, instead of only going up in time, were a tremendous knot, and then, when we cut through the knot, by the plane corresponding to a fixed time, we would see many, many world lines and that would represent many electrons—except for one thing. If in one section this is an ordinary electron world line, in the section in which t reversed itself and is coming back from the future we have the wrong sign to the proper time—to the proper four velocities—and that's equivalent to changing the sign of the charge, and, therefore, that part of a path would act like a positron." "But, Professor," I said, "there aren't as many positrons as electrons." "Well, maybe they are hidden in the protons or something," he said. I did not take the idea that all the electrons were the same one from him as seriously as I took the observation that positrons could simply be represented as electrons going from the future to the past in a back section of their world lines. That, I stole!

To summarize, when I was done with this, as a physicist I had gained two things. One, I knew many different ways of formulating classical electrodynamics, with many different mathematical forms. I got to know how to express the subject every which way. Second, I had a point of view—the overall space-time point of view—and a disrespect for the Hamiltonian method of describing physics.

I would like to interrupt here to make a remark. The fact that electrodynamics can be written in so many ways—the differential equations of Maxwell, various minimum principles with fields, minimum principles without fields, all different kinds of ways—was something I knew but have never understood. It always seems odd to me that the fundamental laws of physics, when discovered, can appear in so many different forms that are not apparently identical at first, but, with a little mathematical fiddling you can show the relationship. An example of that is the Schrödinger equation and the Heisenberg formulation of quantum mechanics. I don't know why this is—it remains a mystery, but it was something I learned from experience. There is always another way to say the same thing that doesn't look at all like the way you said it before. I don't know what the reason for this is. I think it is somehow a representation of the simplicity of nature. A thing like the inverse square law is just right to be represented by the solution of Poisson's equation, which, therefore, is a very different way to say the same thing that doesn't look at all like the way you said it before. I don't know what it means, that nature chooses these curious forms, but maybe that is a way of defining simplicity. Perhaps a thing is simple if you can describe it fully in several different ways without immediately knowing that you are describing the same thing.

I was now convinced that since we had solved the problem of classical electrodynamics (and completely in accordance with my program from M.I.T., with only direct interaction

between particles, in a way that made fields unnecessary) everything was definitely going to be all right. I was convinced that all I had to do was make a quantum theory analogous to the classical one and everything would be solved.

So, the problem is only to make a quantum theory which has as its classical analog this expression 1. Now, there is no unique way to make a quantum theory from classical mechanics, although all the textbooks make believe there is. What they would tell you to do was find the momentum variables and replace them by $(\hbar/i)\ (\partial/\partial x)$, but I couldn't find a momentum variable, as there wasn't any.

The character of quantum mechanics of the day was to write things in the famous Hamiltonian way—in the form of a differential equation, which described how the wave function changes from instant to instant, and in terms of an operator, H. If the classical physics could be reduced to a Hamiltonian form, everything was all right. Now, least action does not imply a Hamiltonian form if the action is a function of anything more than positions and velocities at the same moment. If the action is of the form of the integral of a function (usually called the Lagrangian) of the velocities and positions at the same time

$$S = \int L(\dot{x}, x)dt \tag{2}$$

then you can start with the Lagrangian and then create a Hamiltonian and work out the quantum mechanics, more or less uniquely. But this expression 1 involves the key variables, positions, at two different times and therefore it was not obvious what to do to make the quantum mechanical analog.

I tried—I would struggle in various ways. One of them was this. If I had harmonic oscillators interacting with a delay in time, I could work out what the normal modes were and guess that the quantum theory of the normal modes was the same as for simple oscillators and kind of work my way back in terms of the original variables. I succeeded in doing that, but I hoped then to generalize to other than a harmonic oscillator, but I learned to my regret something which many people have learned. The harmonic oscillator is too simple; very often you can work out what it should do in quantum theory without getting much of a clue as to how to generalize your results to other systems.

So that didn't help me very much, but when I was struggling with this problem, I went to a beer party in the Nassau Tavern in Princeton. There was a gentleman, newly arrived from Europe (Herbert Jehle) who came and sat next to me. Europeans are much more serious than we are in America because they think that a good place to discuss intellectual matters is a beer party. So, he sat by me and asked, "what are you doing" and so on, and I said, "I'm drinking beer." Then I realized that he wanted to know what work I was doing and I told him I was struggling with this problem, and I simply turned to him and said, "listen, do you know any way of doing quantum mechanics, starting with action—where the action integral comes into the quantum mechanics?" "No," he said, "but Dirac has a paper in which the Lagrangian, at least, comes into quantum mechanics. I will show it to you tomorrow."

Next day we went to the Princeton Library; they have little rooms on the side to discuss things, and he showed me this paper. What Dirac said was the following: There is in quantum mechanics a very important quantity which carries the wave function from one time to another, besides the differential equation but equivalent to it, a kind of a kernel, which we might call $K(x',x)$, which carries the wave function $\psi(x)$ known at time t, to the wave function $\psi(x')$ at time $t+\epsilon$. Dirac points out that this function K was *analogous* to the quantity in classical mechanics that you would calculate if you took the exponential of $i\epsilon$, multiplied by the Lagrangian $L(x, x)$, imagining that these two positions x, x' corresponded to t and $t+\epsilon$. In other words,

$$K(x', x) \text{ is analogous to } e^{i\epsilon L\left(\frac{x'-x}{\epsilon}, x\right)}.$$

Professor Jehle showed me this, I read it, he explained it to me, and I said, "what does he mean, they are analogous; what does that mean, *analogous*? What is the use of that?" He said, "you Americans! You always want to find a use for everything!" I said that I thought that Dirac must mean that they were equal. "No," he explained, "he doesn't mean they are equal." "Well," I said, "let's see what happens if we make them equal."

So, I simply put them equal, taking the simplest example where the Lagrangian is ½ $M\dot{x}^2 - V(x)$ but soon found I had to put a constant of proportionality A in, suitably adjusted. When I substituted $Ae^{i\epsilon L}$ for K to get

$$\psi(x', t+\epsilon) = \int A \exp \left[\frac{i\epsilon}{\hbar} L\left(\frac{x'-x}{\epsilon}, x\right)\right] \psi(x, t)\, dx \tag{3}$$

and just calculated things out by Taylor series expansion, out came the Schrödinger equation. So, I turned to Professor Jehle, not really understanding, and said, "well, you see Professor Dirac meant that they were proportional." Professor Jehle's eyes were bugging out—he had taken out a little notebook and was rapidly copying it down from the blackboard, and said, "no, no, this is an important discovery. You Americans are always trying to find out how something can be used. That's a good way to discover things!" So, I thought I was finding out what Dirac meant, but, as a matter of fact, I had made the discovery that what Dirac thought was analogous was, in fact, equal. I had then, at least, the connection between the Lagrangian and quantum mechanics, but still with wave functions and infinitesimal times.

It must have been a day or so later, when I was lying in bed thinking about these things, that I imagined what would happen if I wanted to calculate the wave function at a finite time interval later.

I would put one of these factors $e^{i\epsilon L}$ in here, and that would give me the wave functions the next moment, $t+\epsilon$, and then I could substitute that back into 3 to get another factor of $e^{i\epsilon L}$ and get the wave function the next moment, $t + 2\epsilon$, and so on and so on. In that way I found myself thinking of a large number of integrals, one after the other in sequence. In the integrand was the product of the exponentials, which, of course, was the exponential of the sum of terms like ϵL. Now, L is the Lagrangian and ϵ is like the time interval dt, so that if you took a sum of such terms, that's exactly like an integral. That's like Riemann's formula for the integral $\int L dt$; you just take the value of each point and add them together. We are to take the limit as ϵ—0, of course. Therefore, the connection between the wave function of one instant and the wave function of another instant a

finite time later could be obtained by an infinite number of integrals (because ϵ goes to zero, of course) of exponential $(iS/\hbar)$ where S is the action expression 2. At last, I had succeeded in representing quantum mechanics directly in terms of the action S.

This led later on to the idea of the amplitude for a path—that for each possible way that the particle can go from one point to another in space-time, there's an amplitude. That amplitude is e to the $i/\hbar$ times the action for the path. Amplitudes from various paths superpose by addition. This then is another, a third, way of describing quantum mechanics, which looks quite different than that of Schrödinger or Heisenberg, but which is equivalent to them.

Now immediately after making a few checks on this thing, what I wanted to do, of course, was to substitute the action 1 for the other, 2. The first trouble was that I could not get the thing to work with the relativistic case of spin one-half. However, although I could deal with the matter only non-relativistically, I could deal with the light or the photon interactions perfectly well by just putting the interaction terms of 1 into any action, replacing the mass terms by the non-relativistic $(M\dot{x}^2/2)\ dt$. When the action had a delay, as it now had, and involved more than one time, I had to lose the idea of a wave function. That is, I could no longer describe the program as, given the amplitude for all positions at a certain time, to compute the amplitude at another time. However, that didn't cause very much trouble. It just meant developing a new idea. Instead of wave functions we could talk about this: that if a source of a certain kind emits a particle, and a detector is there to receive it, we can give the amplitude that the source will emit and the detector receive. We do this without specifying the exact instant that the source emits or the exact instant that any detector receives, without trying to specify the state of anything at any particular time in between, but by just finding the amplitude for the complete experiment. And, then we could discuss how that amplitude would change if you had a scattering sample in between, as you rotated and changed angles, and so on, without really having any wave functions.

It was also possible to discover what the old concepts of energy and momentum would mean with this generalized action. And so I believed that I had a quantum theory of classical electrodynamics—or rather of this new classical electrodynamics described by action 1. I made a number of checks. If I took the Frenkel field point of view, which you remember was more differential, I could convert it directly to quantum mechanics in a more conventional way. The only problem was how to specify in quantum mechanics the classical boundary conditions to use only half-advanced and half-retarded solutions. By some ingenuity in defining what that meant, I found that the quantum mechanics with Frenkel fields, plus a special boundary condition, gave me back this action 1, in the new form of quantum mechanics with a delay. So, various things indicated that there wasn't any doubt I had everything straightened out.

It was also easy to guess how to modify the electrodynamics, if anybody ever wanted to modify it. I just changed the delta to an f, just as I would for the classical case. So, it was very easy, a simple thing. To describe the old retarded theory without explicit mention of fields I would have to write probabilities, not just amplitudes. I would have to square my amplitudes and that would involve double path integrals in which there are two S's and so forth. Yet, as I worked out many of these things and studied different forms and different boundary conditions, I got a kind of funny feeling that things weren't exactly right. I could not clearly identify the difficulty and in one of the short periods during which I imagined I had laid it to rest, I published a thesis and received my Ph.D.

During the war, I didn't have time to work on these things very extensively, but wandered about on buses and so forth, with little pieces of paper, and struggled to work on it and discovered indeed that there was something wrong, something terribly wrong. I found that if one generalized the action from the nice Lagrangian forms, 2, to these forms, 1, then the quantities which I defined as energy, and so on, would be complex. The energy values of stationary states wouldn't be real and probabilities of events wouldn't add up to 100%. That is, if you took the probability that this would happen and that would happen—everything you could think of would happen—it would not add up to one.

Another problem on which I struggled very hard was to represent relativistic electrons with this new quantum mechanics. I wanted to do it a unique and different way—and not just by copying the operators of Dirac into some kind of an expression and using some kind of Dirac algebra instead of ordinary complex numbers. I was very much encouraged by the fact that in one space dimension I did find a way of giving an amplitude to every path by limiting myself to paths which only went back and forth at the speed of light. The amplitude was simple $(i\epsilon)$ to a power equal to the number of velocity reversals where I have divided the time into steps ϵ and I am allowed to reverse velocity only at such a time. This gives (as ϵ approaches zero) Dirac's equation in two dimensions—one dimension of space and one of time $(\hbar = M = c = 1)$.

Dirac's wave function has four components in four dimensions, but in this case it has only two components, and this rule for the amplitude of a path automatically generates the need for two components. Because if this is the formula for the amplitudes of path, it will not do you any good to know the total amplitude of all paths which come into a given point to find the amplitude to reach the next point. This is because for the next time, if it came in from the right, there is no new factor $i\epsilon$ if it goes out to the right, whereas, if it came in from the left there was a new factor $i\epsilon$. So, to continue this same information forward to the next moment, it was not sufficient information to know the total amplitude to arrive, but you had to know the amplitude to arrive from the right and the amplitude to arrive from the left, independently. If you did, however, you could then compute both of those again independently and thus you had to carry two amplitudes to form a differential equation (first order in time).

And so I dreamed that if I were clever I would find a formula for the amplitude of a path that was beautiful and simple for three dimensions of space and one of time, which would be equivalent to the Dirac equation, and for which the four components, matrices, and all those other mathematical funny things would come out as a simple consequence—I have never succeeded in that either. But, I did want to mention some of the unsuccessful things on which I spent almost as much effort as on the things that did work.

To summarize the situation a few years after the war, I would say I had much experience with quantum electrodynamics, at least in the knowledge of many different ways of formulating it, in terms of path integrals of actions and in other forms. One of the important by-products, for example, of much experience in these simple forms was that it was easy to see how to combine together what were in those days called the longitudinal and transverse fields, and in general to see clearly the relativistic invariance of the theory. Because of the need to do things differentially there had been, in the standard quantum electrodynamics, a complete split of the field into two parts, one which is called the longitudinal part and the other mediated by the photons, or transverse waves. The longitudinal part was described by a Coulomb potential acting instantaneously in the Schrödinger equation, while the transverse part had an entirely different description in terms of quantization of the transverse waves. This separation depended upon the relativistic tilt of your axes in space-time. People moving at different velocities would separate the same field into longitudinal and transverse fields in a different way. Furthermore, the entire formulation of quantum mechanics, insisting, as it did, on the wave function at a given time, was hard to analyze relativistically. Somebody else in a different coordinate system would calculate the succession of events in terms of wave functions on differently cut slices of space-time and with a different separation of longitudinal and transverse parts. The Hamiltonian theory did not look relativistically invariant, although, of course, it was. One of the great advantages of the overall point of view was that you could see the relativistic invariance right away—or, as Schwinger would say, the covariance was manifest. I had the advantage, therefore, of having a manifestedly covariant form for quantum electrodynamics with suggestions for modifications and so on. I had the disadvantage that if I took it too seriously—I mean, if I took it seriously at all in this form—I got into trouble with these complex energies and the failure of adding probabilities to one and so on. I was unsuccessfully struggling with that.

Then Lamb did his experiment, measuring the separation of the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ levels of hydrogen, finding it to be about 1000 megacycles of frequency difference. Professor Bethe, with whom I was then associated at Cornell, is a man who has this characteristic: If there's a good experimental number you've got to figure it out from theory. So, he forced the quantum electrodynamics of the day to give him an answer to the separation of these two levels. He pointed out that the self-energy of an electron itself is infinite, so that the calculated energy of a bound electron should also come out infinite. But, when you calculated the separation of the two energy levels in terms of the corrected mass instead of the old mass, it would turn out, he thought, that the theory would give convergent finite answers. He made an estimate of the splitting that way and found out that it was still divergent, but he guessed that was probably due to the fact that he used an unrelativistic theory of the matter. Assuming it would be convergent if relativistically treated, he estimated he would get about a thousand megacycles for the Lamb-shift, and thus, made the most important discovery in the history of the theory of quantum electrodynamics. He worked this out on the train from Ithaca, New York, to Schenectady and telephoned me excitedly from Schenectady to tell me the result, which I don't remember fully appreciating at the time.

Returning to Cornell, he gave a lecture on the subject, which I attended. He explained that it gets very confusing to figure out exactly which infinite term corresponds to what in trying to make the correction for the infinite change in mass. If there were any modifications whatever, he said, even though not physically correct (that is, not necessarily the way nature actually works) but any modification whatever at high frequencies, which would make this correction finite, then there would be no problem at all to figuring out how to keep track of everything. You just calculate the finite mass correction Δm to the electron mass m_o, substitute the numerical values of $m_o+\Delta m$ for m in the results for any other problem and all these ambiguities would be resolved. If, in addition, this method were relativistically invariant, then we would be absolutely sure how to do it without destroying relativistic invariance.

After the lecture, I went up to him and told him, "I can do that for you. I'll bring it in for you tomorrow." I guess I knew every way to modify quantum electrodynamics known to man, at the time. So, I went in next day, and explained what would correspond to the modification of the delta-function to f and asked him to explain to me how you calculate the self-energy of an electron, for instance, so we can figure out if it's finite.

I want you to see an interesting point. I did not take the advice of Professor Jehle to find out how it was useful. I never used all that machinery which I had cooked up to solve a single relativistic problem. I hadn't even calculated the self-energy of an electron up to that moment, and was studying the difficulties with the conservation of probability, and so on, without actually doing anything, except discussing the general properties of the theory.

But now I went to Professor Bethe, who explained to me on the blackboard, as we worked together, how to calculate the self-energy of an electron. Up to that time when you did the integrals they had been logarithmically divergent. I told him how to make the relativistically invariant modifications that I thought would make everything all right. We set up the integral which then diverged at the sixth power of the frequency instead of logarithmically!

So, I went back to my room and worried about this thing and went around in circles trying to figure out what was wrong because I was sure physically everything had to come out finite. I couldn't understand how it came out infinite. I became more and more interested and finally realized I had to learn how to make a calculation. So, ultimately, I taught myself how to calculate the self-energy of an electron, working my patient way through the terrible confusion of those days of negative energy states and holes and longitudinal contributions and so on. When I finally found out how to do it and did it with the modifications I wanted to suggest, it turned out that it was nicely convergent and finite, just as I had expected. Professor Bethe and I have never been able to discover what we did wrong on that blackboard two months before, but apparently we just went off somewhere and we have never been able to figure out where. It turned out that what I had proposed, if we had carried it out without making a mistake, would have been all right and would

have given a finite correction. Anyway, it forced me to go back over all this and to convince myself physically that nothing can go wrong. At any rate, the correction to mass was now finite, proportional to $ln(ma/\hbar)$ where a is the width of that function f which was substituted for δ. If you wanted an unmodified electrodynamics, you would have to take a equal to zero, getting an infinite mass correction. But, that wasn't the point. Keeping a finite, I simply followed the program outlined by Professor Bethe and showed how to calculate all the various things—the scatterings of electrons from atoms without radiation, the shifts of levels and so forth—calculating everything in terms of the experimental mass, and noting that the results, as Bethe suggested, were not sensitive to a in this form and even had a definite limit as $a \rightarrow 0$.

The rest of my work was simply to improve the techniques then available for calculations, making diagrams to help analyze perturbation theory quicker. Most of this was first worked out by guessing—you see, I didn't have the relativistic theory of matter. For example, it seemed to me obvious that the velocities in non-relativistic formulas have to be replaced by Dirac's matrix a or in the more relativistic forms by the operators γ_μ. I just took my guesses from the forms that I had worked out using path integrals for non-relativistic matter, but relativistic light. It was easy to develop rules of what to substitute to get the relativistic case. I was very surprised to discover that it was not known at that time that every one of the formulas that had been worked out so patiently by separating longitudinal and transverse waves could be obtained from the formula for the transverse waves alone, if instead of summing over only the two perpendicular polarization directions you would sum over all four possible directions of polarization. It was so obvious from the action 1 that I thought it was general knowledge and would do it all the time. I would get into arguments with people, because I didn't realize they didn't know that; but, it turned out that all their patient work with the longitudinal waves was always equivalent to just extending the sum on the two transverse directions of polarization over all four directions. This was one of the amusing advantages of the method. In addition, I included diagrams for the various terms of the perturbation series, improved notations to be used, worked out easy ways to evaluate integrals, which occurred in these problems, and so on, and made a kind of handbook on how to do quantum electrodynamics.

But one step of importance that was physically new was involved with the negative energy sea of Dirac, which caused me so much logical difficulty. I got so confused that I remembered Wheeler's old idea about the positron being, maybe, the electron going backward in time. Therefore, in the time-dependent perturbation theory that was usual for getting self-energy, I simply supposed that for a while we could go backward in the time, and looked at what terms I got by running the time variables backward. They were the same as the terms that other people got when they did the problem a more complicated way, using holes in the sea, except, possibly, for some signs. These I at first determined empirically by inventing and trying some rules.

I have tried to explain that all the improvements of relativisitc theory were at first more or less straightforward, semi-empirical shenanigans. Each time I would discover something, however, I would go back and I would check it so many ways, compare it to every problem that had been done previously in electrodynamics (and later, in weak coupling meson theory) to see if it would always agree, and so on, until I was absolutely convinced of the truth of the various rules and regulations which I concocted to simplify all the work.

During this time, people had been developing meson theory, a subject I had not studied in any detail. I became interested in the possible application of my methods to perturbation calculations in meson theory. But, what was meson theory? All I knew was that meson theory was something analogous to electrodynamics, except that particles corresponding to the photon had a mass. It was easy to guess that the δ-function in 1, which was a solution of d'Alembertian equals zero, was to be changed to the corresponding solution of d'Alembertian equals m^2. Next, there were different kinds of mesons—the ones in closest analogy to photons, coupled via $\gamma_\mu\gamma_\mu$, are called vector mesons; there were also scalar mesons. Well, maybe that corresponds to putting unity in place of the γ_μ, perhaps what they called "pseudo vector coupling," and I would guess what that probably was. I didn't have the knowledge to understand the way these were defined in the conventional papers because they were expressed at that time in terms of creation and annihilation operators, and so on, which I had not successfully learned. I remember that when someone had started to teach me about creation and annihilation operators, that this operator creates an electron, I said, "how do you create an electron? It disagrees with the conservation of charge," and in that way I blocked my mind from learning a very practical scheme of calculation. Therefore, I had to find as many opportunities as possible to test whether I guessed right as to what the various theories were.

One day a dispute arose at a Physical Society meeting as to the correctness of a calculation by Slotnick of the interaction of an electron with a neutron, using pseudo scalar theory with pseudo vector coupling and also pseudo scalar theory with pseudo scalar coupling. He had found that the answers were not the same; in fact, by one theory, the result was divergent, although convergent with the other. Some people believed that the two theories must give the same answer for the problem. This was a welcome opportunity to test my guesses as to whether I really did understand what these two couplings were. So, I went home, and during the evening I worked out the electron neutron scattering for the pseudo scalar and pseudo vector coupling, saw they were not equal and subtracted them, and worked out the difference in detail. The next day, at the meeting, I saw Slotnick and said, "Slotnick, I worked it out last night, I wanted to see if I got the same answers you do. I got a different answer for each coupling—but, I would like to check in detail with you because I want to make sure of my methods." And, he said, "what do you mean you worked it out last night, it took me six months!" And, when we compared the answers he looked at mine and he asked, "what is that Q in there, that variable Q?" (I had expressions like ($\tan^{-1}Q/Q$ etc.). I said, "that's the momentum transferred by the electron, the electron deflected by different angles." "Oh," he said, "no, I only have the limiting value as Q approaches zero; the forward scattering." Well, it was easy enough to just substitute Q equals zero in my form and I then got the same answers as he

did. But, it took him six months to do the case of zero momentum transfer, whereas, during one evening I had done the finite and arbitrary momentum transfer. That was a thrilling moment for me, like receiving the Nobel Prize, because that convinced me, at last, I did have some kind of method and technique and understood how to do something that other people did not know how to do. That was my moment of triumph in which I realized I really had succeeded in working out something worthwhile.

At this stage, I was urged to publish this because everybody said it looks like an easy way to make calculations, and wanted to know how to do it. I had to publish it, missing two things; one was proof of every statement in a mathematically conventional sense. Often, even in a physicist's sense, I did not have a demonstration of how to get all of these rules and equations from conventional electrodynamics. But, I did know from experience, from fooling around, that everything was, in fact, equivalent to the regular electrodynamics and had partial proofs of many pieces, although I never really sat down, like Euclid did for the geometers of Greece, and made sure that you could get it all from a single simple set of axioms. As a result, the work was criticized, I don't know whether favorably or unfavorably, and the "method" was called the "intuitive method." For those who do not realize it, however, I should like to emphasize that there is a lot of work involved in using this "intuitive method" successfully. Because no simple clear proof of the formula or idea presents itself, it is necessary to do an unusually great amount of checking and rechecking for consistency and correctness in terms of what is known, by comparing to other analogous examples, limiting cases, etc. In the face of the lack of direct mathematical demonstration, one must be careful and thorough to make sure of the point, and one should make a perpetual attempt to demonstrate as much of the formula as possible. Nevertheless, a very great deal more truth can become known than can be proven.

It must be clearly understood that in all this work I was representing the conventional electrodynamics with retarded interaction, and not my half-advanced and half-retarded theory corresponding to 1. I merely use 1 to guess at forms. And one of the forms I guessed at corresponded to changing δ to a function f of width a^2, so that I could calculate finite results for all of the problems. This brings me to the second thing that was missing when I published the paper, an unresolved difficulty. With δ replaced by f the calculations would give results which were not "unitary," that is, for which the sum of the probabilities of all alternatives was not unity. The deviation from unity was very small, in practice, if a was very small. In the limit that I took a very tiny, it might not make any difference. And so the process of the renormalization could be made, you could calculate everything in terms of the experimental mass and then take the limit, and the apparent difficulty that the unitary is violated temporarily seems to disappear. I was unable to demonstrate that, as a matter of fact, it does.

It is lucky that I did not wait to straighten out that point, for as far as I know, nobody has yet been able to resolve this question. Experience with meson theories, with stronger couplings, and with strongly coupled vector photons, although not proving anything, convinces me that if the coupling were stronger, or if you went to a higher order (137th order of perturbation theory for electrodynamics), this difficulty would remain in the limit and there would be real trouble. That is, I believe there is really no satisfactory quantum electrodynamics, but I'm not sure. And I believe that one of the reasons for the slowness of present day progress in understanding the strong interactions is that there isn't any relativistic theoretical model from which you can really calculate everything. Although it is usually said that the difficulty lies in the fact that *strong* interactions are too hard to calculate, I believe it is really because strong interactions in field theory have no solution, have no sense—they're either infinite, or, if you try to modify them, the modification destroys the unitarity. I don't think we have a completely satisfactory relativistic quantum mechanical model, even one that doesn't agree with nature but, at least, agrees with the logic that the sum of probability of all alternatives has to be 100%. Therefore, I think that the renormalization theory is simply a way to sweep the difficulties of the divergences of electrodynamics under the rug. I am, of course, not sure of that.

This completes the story of the development of the space-time view of quantum electrodynamics. I wonder if anything can be learned from it. I doubt it. It is most striking that most of the ideas developed in the course of this research were not ultimately used in the final result. For example, the half-advanced and half-retarded potential was not finally used, the action expression 1 was not used, the idea that charges do not act on themselves was abandoned. The path integral formulation of quantum mechanics was useful for guessing at final expressions and at formulating the general theory of electrodynamics in new ways—although, strictly it was not absolutely necessary. The same goes for the idea of the positron being a backward-moving electron; it was very convenient, but not strictly necessary for the theory because it is exactly equivalent to the negative energy sea point of view.

We are struck by the very large number of different physical viewpoints and widely different mathematical formulations that are all equivalent to one another. The method used here, of reasoning in physical terms, therefore, appears to be extremely inefficient. On looking back over the work, I can only feel a kind of regret for the enormous amount of physical reasoning and mathematical re-expression which ends by merely re-expressing what was previously known, although in a form which is much more efficient for the calculation of specific problems. Would it not have been much easier to simply work entirely in the mathematical framework to elaborate a more efficient expression? This would certainly seem to be the case, but it must be remarked that although the problem actually solved was only such a reformulation, the problem originally tackled was the (possibly still unsolved) problem of avoidance of the infinities of the usual theory. Therefore, a new theory was sought, not just a modification of the old. Although the quest was unsuccessful, we should look at the question of the value of physical ideas in developing a *new* theory.

Many different physical ideas can describe the same physical reality. Thus, classical electrodynamics can be described by a field view, or an action at a distance view, etc. Originally, Maxwell filled space with idler wheels, and Faraday with field lines, but somehow the Maxwell equations themselves are pristine and independent of the elaboration of words attempting a physical description. The only true

physical description is that describing the experimental meaning of the quantities in the equation—or better, the way the equations are to be used in describing experimental observations. This being the case, perhaps the best way to proceed is to try to guess equations, and disregard physical models or descriptions. For example, McCullough guessed the correct equations for light propagation in a crystal long before his colleagues using elastic models could make head or tail of the phenomena, or again, Dirac obtained his equation for the description of the electron by an almost purely mathematical proposition. A simple physical view by which all the contents of this equation can be seen is still lacking.

Therefore, I think equation guessing might be the best method for proceeding to obtain the laws for the part of physics which is presently unknown. Yet, when I was much younger, I tried this equation guessing and I have seen many students try this, but it is very easy to go off in wildly incorrect and impossible directions. I think the problem is not to find the *best* or most efficient method for proceeding to a discovery, but to find any method at all. Physical reasoning does help some people to generate suggestions as to how the unknown may be related to the known. Theories of the known which are described by different physical ideas may be equivalent in all their predictions and hence scientifically indistinguishable. However, they are not psychologically identical when one is trying to move from that base into the unknown. For different views suggest different kinds of modifications which might be made and hence are not equivalent in the hypotheses one generates from them in one's attempt to understand what is not yet understood. I, therefore, think that a good theoretical physicist today might find it useful to have a wide range of physical viewpoints and mathematical expressions of the same theory (for example, of quantum electrodynamics) available to him. This may be asking too much of one man. Then new students should as a class have this. If every individual student follows the same current fashion in expressing and thinking about electrodynamics or field theory, then the variety of hypotheses being generated to understand strong interactions, say, is limited. Perhaps rightly so, for possibly the chance is high that the truth lies in the fashionable direction. But, on the off-chance that it is in another direction —a direction obvious from an unfashionable view of field theory— who will find it? Only someone who has sacrificed himself by teaching himself quantum electrodynamics from a peculiar and unusual point of view, one that he may have to invent for himself. I say sacrificed himself because he most likely will get nothing from it, because the truth may lie in another direction, perhaps even the fashionable one.

But, if my own experience is any guide, the sacrifice is really not great because if the peculiar viewpoint taken is truly experimentally equivalent to the usual in the realm of the known there is always a range of applications and problems in this realm for which the special viewpoint gives one a special power and clarity of thought, which is valuable in itself. Furthermore, in the search for new laws, you always have the psychological excitement of feeling that possibly nobody has yet thought of the crazy possibility you are looking at right now.

So what happened to the old theory that I fell in love with as a youth? Well, I would say it's become an old lady, who has very little that's attractive left in her, and the young today will not have their hearts pound when they look at her anymore. But, we can say the best we can for any old woman, that she has been a very good mother and has given birth to some very good children. And, I thank the Swedish Academy of Sciences for complimenting one of them. Thank you.

Mathematics can help physics, but they are two quite different activities.

25 The Relation of Mathematics to Physics

Richard P. Feynman

Excerpt from his book, *The Character of Physical Law*, published in 1965.

I should like to say a few things on the relation of mathematics and physics which are a little more general. Mathematicians are only dealing with the structure of reasoning, and they do not really care what they are talking about. They do not even need to *know* what they are talking about, or, as they themselves say, whether what they say is true. I will explain that. You state the axioms, such-and-such is so, and such-and-such is so. What then? The logic can be carried out without knowing what the such-and-such words mean. If the statements about the axioms are carefully formulated and complete enough, it is not necessary for the man who is doing the reasoning to have any knowledge of the meaning of the words in order to deduce new conclusions in the same language. If I use the word triangle in one of the axioms there will be a statement about triangles in the conclusion, whereas the man who is doing the reasoning may not know what a triangle is. But I can read his reasoning back and say, 'Triangle, that is just a three-sided what-have-you, which is so-and-so', and then I know his new facts. In other words, mathematicians prepare abstract reasoning ready to be used if you have a set of axioms about the real world. But the physicist has meaning to all his phrases. That is a very important thing that a lot of people who come to physics by way of mathematics do not appreciate. Physics is not mathematics, and mathematics is not physics. One helps the other. But in physics you have to have an understanding of the connection of words with the real world. It is

necessary at the end to translate what you have figured out into English, into the world, into the blocks of copper and glass that you are going to do the experiments with. Only in that way can you find out whether the consequences are true. This is a problem which is not a problem of mathematics at all.

Of course it is obvious that the mathematical reasonings which have been developed are of great power and use for physicists. On the other hand, sometimes the physicists' reasoning is useful for mathematicians.

Mathematicians like to make their reasoning as general as possible. If I say to them, 'I want to talk about ordinary three dimensional space', they say 'If you have a space of n dimensions, then here are the theorems'. 'But I only want the case 3', 'Well, substitute $n = 3$.'! So it turns out that many of the complicated theorems they have are much simpler when adapted to a special case. The physicist is always interested in the special case; he is never interested in the general case. He is talking about something; he is not talking abstractly about anything. He wants to discuss the gravity law in three dimensions; he never wants the arbitrary force case in n dimensions. So a certain amount of reducing is necessary, because the mathematicians have prepared these things for a wide range of problems. This is very useful, and later on it always turns out that the poor physicist has to come back and say, 'Excuse me, when you wanted to tell me about four dimensions . . .'

When you know what it is you are talking about, that some symbols represent forces, others masses, inertia, and so on, then you can use a lot of commonsense, seat-of-the-pants feeling about the world. You have seen various things, and you know more or less how the phenomenon is going to behave. But the poor mathematician translates it into equations, and as the symbols do not mean anything to him he has no guide but precise mathematical rigour and care in the argument. The physicist, who knows more or less how the answer is going to come out, can sort of guess part way, and so go along rather rapidly. The mathematical rigour of great precision is not very useful in physics. But one should not criticize the mathematicians on this score. It is not necessary that just because something would be useful to physics they have to do it that way. They are doing their own job. If you want something else, then you work it out for yourself.

Current emphasis on studies of very small systems and very short time intervals, on the one hand, and large-scale objects of astronomical dimensions, on the other, should lead to increasing interaction and unity between them.

26 Where Do We Go From Here?

Arthur E. Ruark

Article in *Physics Today*, 1969.

Because all science feeds on unsolved problems, it is our privilege, from time to time, to make some forecast of the future. Naturally, the forecaster can do nothing about some great surprise that may come, with sudden force, to change the course of a whole science. Nevertheless, in a well developed science such as physics, one can see some invariant driving forces. There are tides in the affairs of physics that drive us onward without cease. The greatest tide of all appears to be explicit faith in the unity and consistency of natural behavior. This faith implies that parts of our subject that develop in relative isolation will come together to form a broader, more perfect structure.

A very striking feature of our times has been the extension of physical and chemical and biological studies to very small sizes and time intervals. I am talking about our ability to deal with atoms, nuclei and elementary particles. Again, there has been extension of our ability to learn about the large-scale features of this universe—this "bourne of space and time," as Tennyson said. These are intellectual and moral endeavors, in the sense that we have to deal with great uniformities in nature; with creation, evolution and final fate.

Here, my unifying thread of thought will be the increasing interaction between subatomic physics and the physics of the heavens. I shall consider some unsolved problems in these fields. The list is highly selective. I have excluded nearly all the things in the mainstream of current effort, in order to include others that now receive little attention but may be in the mainstream in years to come. Let us proceed, beginning with a few topics in fundamental physics.

THE VERY, VERY SMALL

We all know of the close relation between the relativity theory and the quantum theory. However, there are curiosities connected with this matter. Partly they arise because the field on which the game of quantum theory is played is a classical manifold, the field of space and time, or better spoken, "space–time." Let me indicate how these two theories are connected at their very roots.

Quantum theory is a relativistic theory. The basic papers of Louis de Broglie and of Erwin Schrödinger already showed that the waves belonging to a particle of speed v have a phase speed c^2/v, where c is the speed of light. This formula arises from special relativity; if one uses Newtonian mechanics, a wrong result is obtained.

Special relativity deals with space and time coördinates x and t, so that it is usually considered to be a classical theory; that is to say, a nonquantum theory. This seems to be correct when one considers it as a mathematical

scheme; for there is no mention of Planck's constant h in the axioms set up by Albert Einstein. On the other hand, I do not think it is generally understood that this point of view has to be modified a bit when we take a hard look at the *interpretation* of the theory.

In order to use the theory in physics, we have to say what the quantities Δx and Δt stand for, and Einstein made the choice that is really useful. When he said Δx, he meant a length measured with a real meter stick. He did not mean a hypothetical, nonexistent "rigid ruler," the kind talked about in geometry classes. When he said Δt, he meant a time measured with a laboratory clock. Now, this has consequences. The object to be measured is a dynamic thing, and so is the standard. The meter stick is a group of crystals, a vibrating body held together by quantum forces, and so is the clock. This consideration is dramatized somewhat in figure 1. It looks as though we are caught in a vicious circle; we want to study the interiors of atoms with the aid of laboratory standards, and Lo! The standards are made out of the very things we want to study.

True enough, we do not actually thrust a meter stick down into the atom. We have none with divisions fine enough, and we know that such a disturbance of the atom would not be pertinent if we could do so. Actually, we have to study the wavelengths of light emitted (and other useful quantities), recording them always with the aid of gross apparatus—a favorite topic of Niels Bohr.

Always there are experimental troubles. Fundamental ones are shown in figures 2 and 3. Always, we are making use of a chain of experimental results and interpretation, concerned with the whole coupled apparatus and based on special relativity and quantum theory together. A central question is whether we wish to use our ordinary ideas about lengths and distances when we get into the domain of the very, very small; is this practice really bad? Not at all. The physicist is always trying to extend the scope of his laws or to find their limitations. He is a great fellow for cutting Gordian knots; so he says:

"I shall continue to use special relativity and quantum theory as a strange pair of partners, to interpret results of my experiments on collisions between elementary particles; and I shall find out whether I run into discrepancies."

Breakdown?

Nowadays, one kind of search for such discrepancies is called experimentation on the breakdown of quantum electrodynamics. It is carried on by studying, for example, collisions be-

After taking bachelor's, master's and doctor's degrees at Johns Hopkins University, Arthur E. Ruark taught at Yale, Pittsburgh, North Carolina and Alabama universities. He joined the Atomic Energy Commission in 1956 as chief of the controlled thermonuclear program and is now senior associate director of the division of research at the AEC.

tween two electrons; one looks at the distribution of scattered electrons to see whether it agrees with predictions from electrodynamics. As of 1968, there was no clear evidence of trouble,[1] down to inferred distances between the collision partners as small as about 1.8×10^{-14} cm.

The question now arises: Could particle theory continue to make use of the customary space–time concept if a breakdown of electrodynamics were found? Let us see. A failure of present-day theory would simply lead to construction of some new formulation, not to a modification of the space–time picture. People would keep that picture. What they want is *consistency* in theoretical talk over the whole range of space–time dimensions, "from zero to infinity." It will be extremely hard to eject the space–time picture from any part of physics. Curvature may be introduced; broader geometries may be invoked, but the continuous manifold will still be there because of the flexibility with which new physical fields can be introduced when experiments appear to suggest their presence.

Weak and infrequent things

The success of Fred Reines and Clyde Cowan[2] in starting up the subject of experimental neutrino physics showed us that studies involving miniscule cross sections can be worth a great deal of effort. There is also the search for gravitational waves. It is heartening to know that Joseph Weber[3] has really excellent apparatus to look for these waves; his laboratory is full of seismographs and the like, for throwing out spurious effects from tides and earthquakes. It is still more heartening to know that he has some events that are difficult to explain by means of terrestrial disturbances.

We should not forget that there *may* be very weak forces in nature, still undiscovered, aside from the gravitational ones. I do not know of any current search for such forces.

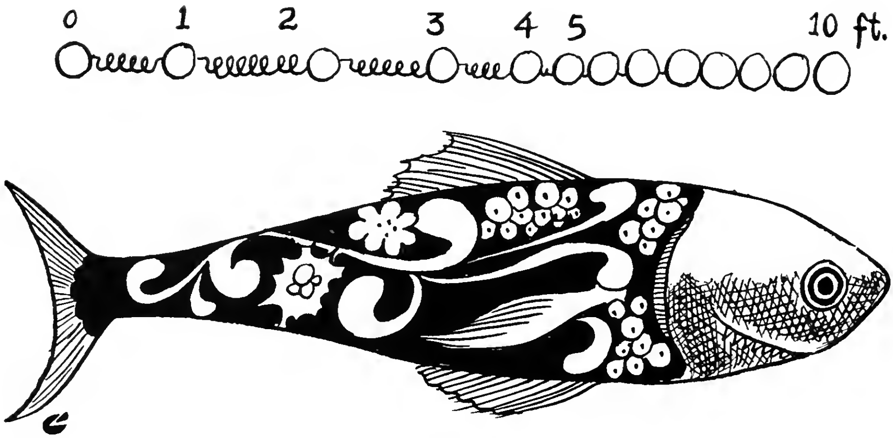
The whole trend in physics has been to assume that particles are extremely well standardized. Nevertheless a few people[4] have been looking for anomalous or nonstandard particles; here I am talking about aberrant electrons, protons, or what-have-you? The resources of modern technique (and in particular, the capabilities of optical spectrographs) are not now being fully used to make some progress with this matter. The trouble is that when one starts to speculate about such particles, the possibilities are very wide; so one must look very selectively for good opportunities to do an interesting experiment.

The search for underlying levels

In recent years we have seen rather extensive searches for an underlying level of simpler things from which a horde of elementary particles might be made. There was the quark search and the search for Dirac magnetic poles; now there is the interest in so-called "W particles." The quark idea, as a mathematical scheme, is indeed ingenious and interesting. The quarks are sometimes thought of as *the* ultimate particles, but there is a trouble with such ideas. If we had quarks, people would just say, "What are they made of?" This is an example of the *Infinite Regression*—a question such that if you answer it you come up against another question of the same kind.

ASTROPHYSICS AND COSMOLOGY

We are all aware of the highly fruitful relations between advances in atomic

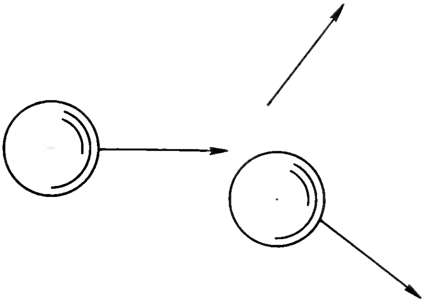


FISHERMAN'S RULE, or how to measure a live fish with a variable rubber Einstein ruler. The fish and the standard are both dynamic objects. —FIG. 1

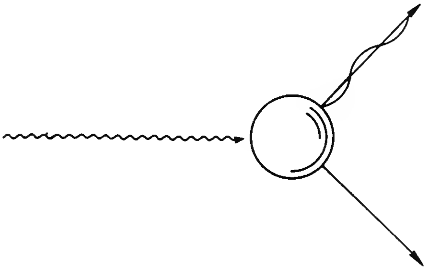
and nuclear physics and those in astrophysics and nebular physics. Furthermore, the fruits of cosmic-ray work, radio astronomy and x-ray astronomy show us that high-energy physics is one essential key to the understanding of very violent astrophysical events.[5] But there is mounting evidence that, in a broader sense, particle physics and cosmology are closely related. Let us turn our attention to a few aspects of this fascinating realm of ideas.

Space–time and matter

It is frequently said that the material content of space and the motion of that material determine the curvature of the space–time manifold. This is often called Mach's principle. Indeed, Einstein's gravitational equations say that a tensor built from curvature quantities is equal to the matter–energy tensor T_{ik}. If T_{ik} is treated as an *arbitrary* source term, the above statement is justified, but we are left with an incomplete story on our hands. Thus, if T_{ik} comes from electromagnetic sources, the fields appearing in it should be taken from Maxwell's equations, written out for curved space–time. Then the curvature and the matter–energy tensor are determined together, from these coupled equations. Einstein proceeded in this way, arriving at his first combined theory of gravitation and electromagnetism. True enough, he abandoned it later for reasons of personal taste, but others have carried on, and this first unified theory is a lively field of research even today, 50 years after it was created. However, a salient question still confronts us. When we proceed to a specific case, that of a single electron for example, do we simply put in the electronic charge as an unexplained parameter? Or do we look for underlying relations whereby the electron can be represented as a curlicue of particular dimensions in space–time? To speak more generally—do we want a completely unified theory of space time and matter, or a dualistic theory? There is a literature on this subject, too extensive for discussion here.[6] An idea of the Mach type runs through it all. If I were asked for a comprehensive generalization of the Mach idea,



ATOMIC BILLIARDS. When we try to measure a coördinate, recoil from the test body alters the coördinates and the momentum under study. —FIG. 2



A PHOTON used for a measurement is affected by its collision with the object under attention. —FIG. 3

I would say, "There is just one manifold. The equations describing physical phenomena contain not only fields defined on that manifold but also quantities characterizing the geometry of the manifold. The connections are such that the fields and the geometrical quantities are determined together, consistently." And I recommend to the reader some interesting studies of a generalized Mach principle, by Mendel Sachs.[7]

This is a good place to ask, "How is it that space has three dimensions?" This question is at least 70 years old. I have seen nothing on the subject that is more than a plausibility argument, but I have a small suggestion as to a fresh approach. Suppose we use the methods of tensor and spinor calculus to examine physical equations in space–time of several dimensions, from two up to six, for example. Let us cover both classical theory and quantum theory, remembering to look closely at the properties of simple solutions that represent point particles; we search for features that appear particularly desirable or unique (or both), in the case of four-dimensional space–time. If such features emerge, we may understand a little better the preference for three space dimensions in this universe. The results would still be plausibility arguments, but if they looked attractive, we would promote them to the status of assumptions; and that would be that.

Consistency: a desirable feature

Perhaps the most significant fact that has emerged from exploration of the distant galaxies is the *general consistency of physical law over very large spaces and long time intervals.* Apparently we are *not* dealing with different bodies of law, linked together only by very weak connections. We appear to be living in a Universe—not in some sort of Diverse, or Polyverse. A cardinal piece of support for this welcome notion is the red shift of Vesto Slipher, Edwin Hubble and Milton Humason. To an approximation, the light from distant galaxies is shifted toward the red, by amounts that can be explained by assuming that they move outward with speeds v, proportional to their distances R from us; the relation is

$$v = 75R,$$

with v in kilometers per second and R in megaparsecs; one megaparsec is 3.09×10^{24} cm.

Allowing for this red shift, we see the same spectral series, the same atomic behavior, that is found here on earth. Of course, this probing out to great distances means that one is looking back a long way in time. What is the inner meaning of this consistency? The distant atoms would not show the spectral series properly if they did not obey the Pauli principle. Those atoms are testifying to identity of the electrons and identity of the nuclei in the whole region available for observation. They are revealing *a most extraordinary degree of quality control in the creation and maintenance* of these particles. Why, not even Rolls-Royce . . . !

Is this uniformity of particle properties due to a uniformity in the properties of space–time itself? Or are these two ideas just the same idea, clothed in different words? I leave the answer to you—or your grandchildren.

Long ago and far away

There is another important fact that bears on the question of universal consistency. Suppose an atom in a galaxy 10^9 light years away emits a parcel of energy characterized by a far-ultraviolet wavelength. Looking aside from experimental difficulties, we can set up a suitable bulb containing sodium vapor, here in our solar system, to receive the light. After 10^9 years an electron may be kicked out of a single atom in that vapor. *If* we believe that an electromagnetic field traveled all that time through empty, darksome space, then we have to say that the field causes a definite amount of energy to appear at a target only 10^{-8} cm in diameter, after running through a distance of about 10^{27} centimeters. Also, from the observed conservation of energy in such processes, we have to conclude that the field does nothing elsewhere.

What shall we say about this result? An orthodox quantum theorist might say, "It is all a matter of chance; this matter was explained in 1927." A thoroughgoing determinist might say, "This astounding accuracy of aim is evidence of extraordinary quality control." A classical relativist might say, "All point events that are connected by light rays are at the *same spot* in space–time. We are dealing with a sort of contact action. From the standpoint of a being who perceives point events directly and intuitively, there is no problem." We possess considerable flexibility in contemplation of these answers or others like them; for each answer is based on some set of axioms, and axioms are arbitrary indeed. The orthodox quantum theorist will say, "Yes, but look at the fruits of my axioms." And we shall reply, "The *fruits* of your axioms are very great indeed, but a large number of very respectable people are not satisfied with the foundations of your theory."

Permanence: a desirable feature

Let us consider the permanence of gross matter. The customary estimates of universe duration lie a little above 10^{10} years. It happens that Reines and his students have found lower limits for the lifetimes of electrons and nucleons by looking for their decay.[8] There are some nuances, but roughly the half-life figures are: for the electron, more than 2×10^{21} years; for nucleons, more than 10^{27} years. Thus we are confronted with a terrific factor of safety, 10^{11} at least, relative to the universe duration mentioned above. This looks like very good engineering. The stuff is made so it will last.

Diluteness: a convenient feature

People are generally impressed with the vast spaces between the stars of our galaxy, and also the spaces between galaxies, which, on the average, are somewhat like tennis balls 8 meters apart. This diluteness is much to be prized, because violent things happen when big pieces of matter get too close together. I invite your attention to the famous case of the galaxy M 82. A photograph of this galaxy can be found in reference 9. More or less perpendicular to the disk of the galaxy there are great masses of ejected matter, believed to be mostly hydrogen. There was a big explosion in the middle of this galaxy. The products are pouring out at a speed of the order 10^8 cm/sec. It is estimated that this explosion involved disruption of a million stars in the dense core of the galaxy.

Information from far away

How much can we hope to learn about very distant objects? In general, the farther away an object is, the less we can find out about it. Details fuzz out; light signals from the object are fainter; spectra move out to the infrared. It is only in recent times that attention has been paid to the quantitative side of this common observation. Kenneth Metzner and Philip Morrison[10] have calculated the amount of information carried to us by the photons from a distant galaxy in any experiment of limited duration. They consider simple expanding universes of several types. This is a matter worthy of further research, because it can show us the boundary between verifiable physics and unverifiable speculation. Beyond the domains where individual galaxies can be identified—and there are hundreds of millions within sight—there may be others that show up as a faint general background. Astronomers know that they must increase their studies of this faint background light, when more big telescopes come on stream, a few years hence.

If and when they reach the limit of their resources, we shall be confronted with an interesting situation. For a long time philosophers have been saying that physicists continually work on the soluble problems, so that metaphysics is necessarily the bin of unsolved ones. Now I shall leave it to the reader to ponder the situation of an experimental science that reaches a limit because the objects under investigation cannot provide sufficient amounts of information to our detectors to give the answers we should like to know.

EPILOGUE

I have pointed out some lines of endeavor that lie at or beyond the present limits of our capabilities, and I have only two hints for those who may choose to attack these matters. The first is that one should pay close attention to a method used by Renè Descartes. I call it the "Method of Complete Skepticism." He adopted a systematic policy of denying any statement he was considering and of looking at the consequences. The second hint is connected with economy and simplicity of thought. I quote the famous dictum of William of Occam: "Entia non multiplicanda sunt, praeter necessitatem." Entities are not to be multiplied except for reasons of necessity.

In closing, I mention once more the consistency, the connectivity, revealed by physical studies up to the present.

Though each of us usually thinks of himself as a part of the universe, this is a one-sided view, for great portions of our surroundings are always exerting their influence upon us. As an overstatement, one might say that the universe is a part of every man. Sir George Thomson[11] says in his book, *The Foreseeable Future*:

> "The universe that includes our perceptions and our feelings is one, and no single part can be put into a ring-fence completely isolated from all the rest."

Therefore I end this story with the thought: The universe is the proper study of mankind.

References

1. W. C. Barber, B. Gittelman, G. K. O'Neill, B. Richter, Phys. Rev. Lett. **16**, 1127 (1966).
2. F. Reines, C. L. Cowan Jr, PHYSICS TODAY **10**, no. 8, 12 (1957).
3. J. Weber, Phys Rev. Lett. **20**, 1307 (1968).
4. G. M. Kukavadze, L. Ya. Memelova, L. Ya. Suvorov, Sov. Phys.–JETP **22**, 272 (1965); E. Fischbach, T. Kirsten, G. A. Schaeffer, Phys. Rev. Lett. **20**, 1012 (1965).
5. S. Colgate, PHYSICS TODAY **22**, no. 1, 27 (1969).
6. J. A. Wheeler, *Geometrodynamics*, Academic Press, New York (1962). D. K. Sen, *Fields and/or Particles*, Academic Press, New York (1968).
7. M. Sachs, PHYSICS TODAY **22**, no. 2, 51 (1969).
8. M. K. Moe, F. Reines, Phys Rev. **140**, B992 (1965); W. R. Kropp Jr, F. Reines, Phys. Rev. **137**, B740 (1965); C. C. Giamati, F. Reines, Phys. Rev. **126**, 2178 (1962).
9. G. F. Burbidge, E. M. Burbidge, A. M. Sandage, Rev. Mod. Phys. **35**, 947 (1963).
10. A. W. K. Metzner P. Morrison, Mon. Not. Roy. Astron. Soc. **119**, 657 (1959).
11. G. P. Thomson, *The Foreseeable Future*, 2nd ed., Viking Press, New York (1960). □

JEREMY BERNSTEIN

Jeremy Bernstein, born in 1929 in Rochester, New York, is Professor of Physics at Stevens Institute of Technology in New Jersey. He was educated at Columbia Grammar School in New York City and received a bachelor's and master's degree in mathematics, and a doctorate in physics from Harvard University. He has done research at the Harvard Cyclotron Laboratory, the Institute for Advanced Study at Princeton, Los Alamos, at the Brookhaven National Laboratories, and is frequently a visiting physicist at CERN (Conseil Européen pour la Recherche Nucléaire) in Geneva. Bernstein is the author of The Analytical Engine: Computers, Past, Present, and Future, Ascent, an account of mountaineering in the Alps, and has written book reviews and Profile articles for the magazine, The New Yorker.

HARRISON SCOTT BROWN

Harrison Scott Brown, born in Sheridan, Wyoming, in 1917, is Professor of Geochemistry at California Institute of Technology and Foreign Secretary of the National Academy of Sciences. He received a B.S. from the University of California and a Ph.D. from Johns Hopkins University. Brown is an editor at large for The Saturday Review and has written The Challenge of Man's Future and Must Destruction Be Our Destiny? His research interests include mass spectroscopy, meteoritics, planet structure a and planetary chemistry.

SIR JAMES CHADWICK

Sir James Chadwick was born in 1891 in Manchester, England; he attended Victoria University there, and then Cambridge University. At the age of eighteen he met Ernest Rutherford with whom he later collaborated in experimental work. Chadwick discovered the neutron in 1932 and for this was awarded the Nobel Prize in Physics in 1935. During World War II he worked for "Tube Alloys," the British equivalent of the Manhattan Project.

OWEN CHAMBERLAIN

Owen Chamberlain, Professor of Physics at the University of California at Berkeley, and Nobel Prize winner in 1959 with Emilio Segré for their demonstration of the existence of the antiproton, was born in San Francisco in 1920. He received his bachelor's degree from Dartmouth College and his Ph.D. from the University of Chicago. During World War II he worked on the Manhattan Project as a civilian physicist. He has been active in civil liberties activities. Some of his special interests in physics are fission, alpha-particle decay, and neutron diffraction in liquids.

LAURA FERMI

Laura Fermi was born in Rome, Italy, in 1907, and studied at the University of Rome. She met Enrico Fermi when she was sixteen; they were married five years later. She has two children. When the anti-Semitic laws appeared in Italy in 1938, the Fermis left for the United States, immediately after he received the Nobel Prize that December. In 1955 she attended the International Conference on the Peaceful Uses of Atomic Energy as historian for the United States and wrote Atoms for the World. She is also author of Atoms in the Family: My Life with Enrico Fermi, and the monographic study, Mussolini.

RICHARD PHILLIPS FEYNMAN

Richard Feynman was born in New York in 1918, and graduated from the Massachusetts Institute of Technology in 1939. He received his doctorate in theoretical physics from Princeton in 1942, and worked at Los Alamos during the Second World War. From 1945 to 1951 he taught at Cornell, and since 1951 has been Tolman Professor of Physics at the California Institute of Technology. Professor Feynman received the Albert Einstein Award in 1954, and in 1965 was named a Foreign Member of the Royal Society. In 1966 he was awarded the Nobel Prize in Physics, which he shared with Shinichero Tomonaga and Julian Schwinger, for work in quantum field theory.

KENNETH W. FORD

Kenneth W. Ford was born in 1917 at West Palm Beach, Florida. He did his undergraduate work at Harvard College. His graduate work at Princeton University was interrupted by two years at Los Alamos and at Project Manhattan in Princeton. He worked on a theory of heavy elementary particles at the Imperial College in London, and at the Max Planck Institute in Göttingen, Germany. Before joining the faculty at the University of California, Irvine, as chairman of the Department of Physics, Mr. Ford was Professor of Physics at Brandeis University.

JAMES FRANCK

James Franck was born in Hamburg, Germany, in 1882, and received his Ph.D. from the University of Berlin. He and Gustav Hertz shared the Nobel Prize in 1925 for their studies which supported the new model of the atom just postulated by Bohr. Franck was Professor of Experimental Physics and Director of the Institute for Experimental Physics at the University of Göttingen. When the Nazis gained increasing power, Franck

demonstrated against the racial laws, and in 1933 he and his family moved to the United States. Here he lectured at Johns Hopkins University and later became Professor of Physical Chemistry at the University of Chicago. He died in 1964.

MARTIN GARDNER

Martin Gardner, the editor of the "Mathematical Games" department of the Scientific American, was born in Tulsa, Oklahoma, in 1914. He received a B.A. in philosophy from the University of Chicago in 1936, worked as a publicity writer for the university, and then wrote for the Tulsa Tribune. During World War II he served in the Navy. Martin Gardner has written numerous short stories as well as professional articles for such journals as Scripta Mathematica and Philosophy of Science, and is the author of the best-selling books, The Annotated Alice, Relativity for the Millions, Fads and Fallacies In the Name of Science, as well as two volumes of the Scientific American Book of Mathematical Puzzles and Diversions.

LEOPOLD INFELD

Leopold Infeld, a co-worker with Albert Einstein in general relativity theory, was born in 1898 in Poland. After studying at the Cracow and Berlin Universities, he became a Rockefeller Fellow at Cambridge where he worked with Max Born in electromagnetic theory, and then a member of the Institute for Advanced Study at Princeton. For eleven years he was Professor of Applied Mathematics at the University of Toronto. He then returned to Poland and became Professor of Physics at the University of Warsaw and until his death on 16 January 1968 he was director of the Theoretical Physics Institute at the university. A member of the presidium of the Polish Academy of Science, Infeld conducted research in theoretical physics, especially relativity and quantum theories. Infeld was the author of The New Field Theory, The World in Modern Science, Quest, Albert Einstein, and with Einstein, The Evolution of Physics.

DAVID LOCKHART JUDD

David Lockhart Judd was born in Chehalis, Washington, in 1923. In 1943 he received his A.B. from Whitman College. He then attended California Institute of Technology and received an M.S. in 1947 and a Ph.D. in physics three years later. From 1951 to the present he has been with the Lawrence Radiation Laboratory at Berkeley, since 1965 as head of the Physics Division. He is also senior lecturer in physics at the University of California, Berkeley. His professional interests include accelerator theory, ion optics, plasma and particle physics, and nonlinear mechanics.

RALPH EUGENE LAPP

Ralph Lapp was born in Buffalo, New York, in 1917. He received his B.S. and Ph.D. in physics from the University of Chicago. He was head of the nuclear physics branch, Office of Naval Research, and since 1950 has been director of the Nuclear Science Service. Lapp is the author of many books concerning the social consequences of modern science, including Must We Hide? and The New Priesthood: The Scientific Elite and The Uses of Power. His interests include cosmic radiation, mass spectroscopy and civil defense.

ERNEST ORLANDO LAWRENCE

Ernest Orlando Lawrence (1901–1958) was born in North Dakota. He received his doctorate from Yale University and then joined the faculty of the University of California at Berkeley. By building with his colleagues, M. S. Livingstone and others, the first successful cyclotron, Lawrence solved one of the major experimental problems of the 1920's and 30's in nuclear physics, that of providing controllable beams of high-energy particles. Lawrence built a series of increasingly more powerful cyclotrons. For these accomplishments and for his research on artificial radioactive elements, Lawrence was awarded the Nobel Prize in Physics in 1939. The element lawrencium is named for him.

GERARD KITCHEN O'NEILL

Professor of Physics at Princeton University, O'Neill was born in Brooklyn, New York, in 1927. He received his bachelor's degree from Swarthmore College and his Ph.D. from Cornell University. Between 1954 and 1959 he was a member of a group that designed the three-billion-volt proton synchrotron now being operated jointly by Princeton and the University of Pennsylvania. More recently he has worked on the design of storage rings, experiments in high-energy physics and spark chambers.

V. LAWRENCE PARSEGIAN

V. Lawrence Parsegian studied at M.I.T., Washington University, and New York University, obtaining his Ph.D. in physics in 1948. He has been professor of nuclear science and engineering at Rensselaer Polytechnic Institute since 1954, and holds the distinguished Chair of Rensselaer professorship. In addition to his research activities, he has chaired a curriculum development project to improve college science teaching.

RUDOLF ERNST PEIERLS

Rudolf Ernst Peierls was born in Berlin in 1907 and received degrees from several universities, including a D.Phil. in Theoretical Physics from

the University of Leipzig in 1929 and a D.Sc. from the University of Manchester, England, in 1936. From 1937 to 1963 he was Professor of Mathematical Physics at Birmingham University. During the early years of World War II he worked on the Atomic Energy Project in Birmingham, and then at Los Alamos between 1943–46. Peierls is now Professor of Theoretical Physics at Oxford University and a Fellow of New College, Oxford. He is the author of The Laws of Nature and Quantum Theory of Solids.

ARTHUR C. RUARK

(see page 256)

ERNEST RUTHERFORD

Lord Rutherford (1871–1937) was born in Nelson, South Island, New Zealand. He graduated from Nelson College. At the University of New Zealand he won a scholarship to attend Cambridge University in England where, stimulated by J.J. Thomson, he studied the electrical nature of matter. As Professor of Physics at McGill University in Montreal, he distinguished the identity of Becquerel's radiations into alpha, beta and gamma rays, and proposed (with Soddy) the concepts of radioactive transmutation and isotopes. Returning to England, he continued his research at the University of Manchester. There he conducted his most famous experiments leading in 1911 to his discovery of the nucleus in the atom. He was awarded the Nobel Prize in Chemistry in 1908 for his experiments in radioactivity. Rutherford returned to Cambridge in 1919 as director of the Cavendish Laboratory.

EMILIO SEGRE

Emilio Segre was born in Tivoli, Italy, in 1905 and received his Ph.D. in physics from the University of Rome in 1928. He was a student of Enrico Fermi from 1934 to 1936, and has published a biography, Enrico Fermi, Physicist (1970). Then he became director of the physics laboratory at Palermo, where he and C. Perrier made the discovery of technetium, the first artificially made element. Segrè and his co-workers also were the first to identify the artificial elements of plutonium and astatine. Segrè was awarded the Nobel Prize in Physics in 1959 for his demonstration with Owen Chamberlain of the existence of the antiproton. He is Professor of Physics at the University of California at Berkeley.

CHARLES PERCY SNOW

Charles Percy Snow, Baron of Leicester, was born in 1905 and educated at University College, Leicester and at Christ's College, Cambridge. Although well known as a novelist, especially dealing with the lives and problems of professional men, he has held such diverse positions as chief of scientific personnel for the Ministry of Labour, Civil Service Commissioner, and a Director of the English Electric Co., Ltd. His writings have been widely acclaimed; among his novels are The Search, The New Men, and Corridors of Power. His nonfiction books on science and its consequences include The Two Cultures and The Scientific Revolution and Science and Government.

LEO SZILARD

Leo Szilard was born in Budapest, Hungary, in 1898, and received his doctorate at the University of Berlin. He was at the Clarendon Laboratory in England and the National Defense Research Division at Columbia University before going to the University of Chicago as Professor of Physics. At the time of his death in May 1964, Szilard was a resident fellow at the Salk Institute for Biological Studies in La Jolla, California. Besides nuclear physics, he did research in a variety of fields including mutations and genetics of bacteria and bacterial viruses. Szilard helped to draft and transmit the famous letter from Einstein to Roosevelt which helped to initiate large-scale work on atomic energy in the United States in 1939. His publications include The Voice of the Dolphins. He was deeply involved with groups that aimed at the peaceful application of science and technology, and in political action toward such ends.

ALVIN MARTIN WEINBERG

Alvin Martin Weinberg, Director of the Oak Ridge National Laboratory in Tennessee, was born in 1915 in Illinois. He graduated from the University of Chicago in 1935 and received his doctorate in physics from Chicago in 1939. He has been on the United States visiting scientist team to Russian nuclear installations, the President's Scientific Advisory Board, and has been awarded the Atoms for Peace Award (1960) and the Lawrence Memorial Award. He is a pianist and dedicated tennis player in his spare time.

CLYDE EDWARD WIEGAND

Clyde Edward Wiegand was born in Long Beach, Washington, in 1915 and graduated from Willamette College in Oregon. He was awarded a Ph.D. in physics from the University of California, where he has been a graduate student of Emilio Segrè. During World War II he went with Segrè to work at the Los Alamos Laboratory. Weigand is now with the University of California at its Lawrence Radiation Laboratory. His research interests include nuclear physics, scattering, and cross-section work with high-energy particles.

VICTOR F. WEISSKOPF

(see page 212)

ROBERT R. WILSON

R.R. Wilson, was born in 1914 in Frontier, Wyoming, and now is director of the National Accelerator Laboratory, Batavia, Illinois, and professor of physics at the University of Chicago. He received his training at the University of California and has taught at Princeton, Harvard, Cornell, and Chicago. Since 1947, Mr. Wilson has been involved in the construction of a series of particle accelerators with which to explore the structure of the proton. He has had formal training as a sculptor in the United States and at the Academia Belli Arte in Rome, and continues actively working in this field.

HERMAN YAGODA

Herman Yagoda, chemist as well as physicist, was born in New York City in 1908. He graduated from Cooper Union and received his master's degree from New York University. Yagoda died in 1964. He had been a chemist for the U. S. Customs Laboratory in New York and was at the Air Force Cambridge Research Laboratories where he conducted research in space physics and cosmic radiation. Yagoda was the author of Radioactive Measurements with Nuclear Emulsions.

GALE YOUNG

Gale Young was born in Baroda, Michigan, in 1912. He received a B.S. from the Milwaukee School of Engineering and a B.S. and M.S. from the University of Chicago. He has taught physics at Chicago University and Olivet College in Michigan. Like many physicists, during World War II Young worked on the Manhattan District Project and was the technical director of the Nuclear Development Association. Since 1961 he has been an executive of the United Nuclear Corporation.

THOMAS JOHN YPSILANTIS

Thomas John Ypsilantis was born in Salt Lake City in 1928. He earned his B.Sc. from the University of Utah and his M.A. and Ph.D. from the University of California, Berkeley. He has been on the faculty at Berkeley since 1957 and is now Associate Professor of Physics. Ypsilantis had a Guggenheim Fellowship in 1959–60, and has been a consultant to the Institute of Defense Analysis. His reseach interests include antiproton interactions, proton polorization in scattering, and pion and nucleon interactions.

Holt/Rinehart/Winston